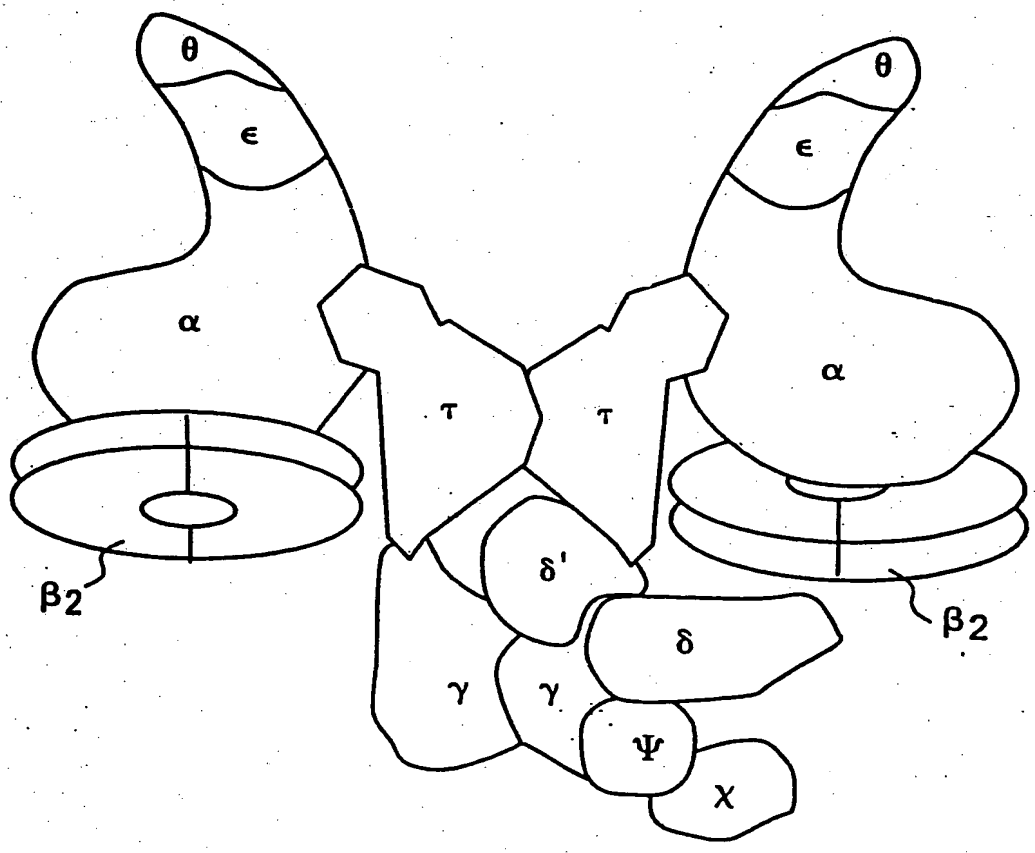
FIG.1



ATP binding

E. coli
MSYQVLARKWRPQTFADVVGQEHVLTALANGLSLGRIH**HAYLFSGTRGVGKT**SIARLLAK
B. subtilis
MSYQALYRVFRPQRFEDVVGQEHITKTLQNALLQKKFS**HAYLFSGPRGTGKT**SAAKIFAK
**** * * *** * *******. .* * * . ******** ** ***** ..**

E. coli
GLNCETGITATPCGVCDNCREIEQGRFVDLIEIDAASRTKVEDTRDLLDNVQYAPARGRF
B. subtilis
AVNCEHAPVDEPCNECAACKGITNGSISDVIEIDAASNNGVDEIRDIRDKVKFAPSAVTY
 .*** **. * *. * .* *.******* . *.. **. * *..**. .

E. coli
KVYLIDEVHMLSRHSFNALL**KTLEEPPEH**VKFLLATTDPQKLPVTILSRCLQFHLKALDV
B. subtilis
KVYIIDEVHMLSIGAFNALL**KTLEEPPEH**CIFILATTEPHKIPLTIISRCQRFDFKRITS
.***** .************** *.****.*.*.*.**.*** .* * .

FIG. 2

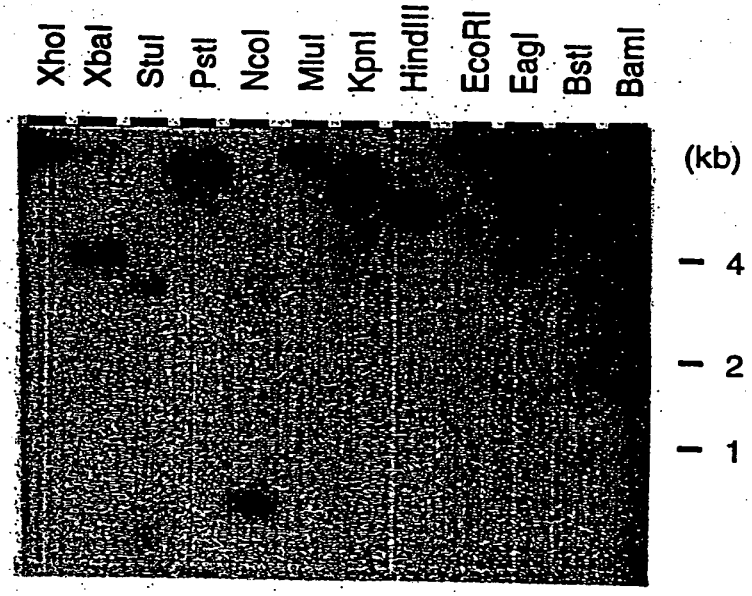

XhoI
XbaI
StuI
PstI
NcoI
MluI
KpnI
HindIII
EcoRI
EagI
BstI
BamI
(kb)
4
2
1

FIG.3

```
TCCGGGGGTG    GGGTTCCCAG    GTAGACCCCG    GCCCCTCCCG    TGAGCCCCTT    TACCCAGGCC    60

GCCACCTCCT    CCAGGGGGGC    CAAGGCGTGC    AAGGAGAGGA    ACGTCCGCAC    CACGCCCTAT    120
                                                                     S.D.
ACTAGCCTT   GTG AGC GCC CTC TAC CGC CGC TTC CGC CCC CTC ACC TTC CAG GAG GTG GTG    180
            met ser ala leu tyr arg arg phe arg pro leu thr phe gln glu val val    (17)

                                                                          CAC
GGG CAG GAG CAC GTG AAG GAG CCC CTC CTC AAG GCC ATC CGG GAG GGG AGG CTC GCC CAG    240
gly gln glu his val lys glu pro leu leu lys ala ile arg glu gly arg leu ala gln    (37)

GCS TAC CTS TTC TCC GGS AC
GCC TAC CTC TTC TCC GGG CCC AGG GGC GTG GGC AAG ACC ACC ACG GCG AGG CTC CTC GCC    300
ala tyr leu phe ser gly pro arg gly val gly lys thr thr thr ala arg leu leu ala    (57)

ATG GCG GTG GGG TGC CAG GGG GAA GAC CCC CCT TGC GGG GTC TGC CCC CAC TGC CAG GCG    360
met ala val gly cys gln gly glu asp pro pro cys gly val cys pro his cys gln ala    (77)

GtG CAG AGG GGC GCC CAC CCG GAC GTG GTG GAC ATT GAC GCC GCC AGC AAC AAC TCC GTG    420
val gln arg gly ala his pro asp val val asp ile asp ala ala ser asn asn ser val    (97)

GAG GAC GTG CGG GAG CTG AGG GAA AGG ATC CAC CTC GCC CCC CTC TCT GCC CCC AGG AAG    480
glu asp val arg glu leu arg glu arg ile his leu ala pro leu ser ala pro arg lys    (117)

                                                                          C
GTC TTC ATC CTG GAC GAG GCC CAC ATG CTC TCC AAA AGC GCC TTC AAC GCC CTC CTC AAG    540
val phe ile leu asp Glu ala his met leu ser lys ser ala phe asn ala leu leu lys    (137)
```

FIG.4A-1

```
TGS CTS CTC CTC GGS GGS CTC GTG
ACC CTG GAG GAG CCC CCG CCC CAC GTC CTC TTC GTC TTC GCC ACC ACC GAG CCC GAG AGG   600
thr leu glu glu pro pro pro his val leu phe val phe ala thr thr glu pro glu arg   (157)

ATG CCC CCC ACC ATC CTC TCC CGC ACC CAG CAC TTC CGC TTC CGC CGC CTC ACG GAG GAG   660
met pro pro thr ile leu ser arg thr gln his phe arg phe arg arg leu thr glu glu   (177)

GAG ATC GCC TTT AAG CTC CGG CGC ATC CTG GAG GCC GTG GGG CGG GAG GCG GAG GAG GAG   720
glu ile ala phe lys leu arg arg ile leu glu ala val gly arg glu ala glu glu glu   (197)

GCC CTC CTC CTC CTC GCC CGC CTG GCG GAC GGG GCC CTT AGG GAC GCG GAA AGC CTC CTG   780
ala leu leu leu leu ala arg leu ala asp gly ala leu arg asp ala glu ser leu leu   (217)

GAG CGC TTC CTC CTC CTG GAA GGC CCC CTC ACC CGG AAG GAG GTG GAG CGC GCC CTA GGC   840
glu arg phe leu leu leu glu gly pro leu thr arg lys glu val glu arg ala leu gly   (237)

TCC CCC CCA GGG ACC GGG GTG GCC GAG ATC GCC GCC TCC CTC GCG AGG GGG AAA ACG GCG   900
ser pro pro gly thr gly val ala glu ile ala ala ser leu ala arg gly lys thr ala   (257)

GAG GCC CTG GGC CTC GCC CGG CGC CTC TAC GGG GAA GGG TAC GCC CCG AGG AGC CTG GTC   960
glu ala leu gly leu ala arg arg leu tyr gly glu gly tyr ala pro arg ser leu val   (277)

TCG GGC CTT TTG GAG GTG TTC CGG GAA GGC CTC TAC GCC GCC TTC GGC CTC GCG GGA ACC  1020
ser gly leu leu glu val phe arg glu gly leu tyr ala ala phe gly leu ala gly thr   (297)

CCC CTT CCC GCC CCG CCC CAG GCC CTG ATC GCC GCC ATG ACC GCC CTG GAC GAG GCC ATG  1080
pro leu pro ala pro pro gln ala leu ile ala ala met thr ala leu asp glu ala met   (317)
```

FIG.4A-2

```
GAG CGC CTC GCC CGC CGC TCC GAC GCC TTA AGC CTG GAG GTG GCC CTC CTG GAG GCG GGA  1140
glu arg leu ala arg arg ser asp ala leu ser leu glu val ala leu leu glu ala gly  (337)

AGG GCC CTG GCC GCC GAG GCC CTA CCC CAG CCC ACG GGC GCT CCT TCC CCA GAG GTC GGC  1200
arg ala leu ala ala glu ala leu pro gln pro thr gly ala pro ser pro glu val gly  (357)

CCC AAG CCG GAA AGC CCC CCG ACC CCG GAA CCC CCA AGG CCC GAG GAG GCG CCC GAC CTG  1260
pro lys pro glu ser pro pro thr pro glu pro pro arg pro glu glu ala pro asp leu  (377)

CGG GAG CGG TGG CGG GCC TTC CTC GAG GCC CTC AGG CCC ACC CTA CGG GCC TTC GTG CGG  1320
arg glu arg trp arg ala phe leu glu ala leu arg pro thr leu arg ala phe val arg  (397)

GAG GCC CGC CCG GAG GTC CGG GAA GGC CAG CTC TGC CTC GCT TTC CCC GAG GAC AAG GCC  1380
glu ala arg pro glu val arg glu gly gln leu cys leu ala phe pro glu asp lys ala  (417)

TTC CAC TAC CGC AAG GCC TCG GAA CAG AAG GTG AGG CTC CTC CCC CTG GCC CAG GCC CAT  1440
phe his tyr arg lys ala ser glu gln lys val arg leu leu pro leu ala gln ala his  (437)

                                                frameshift site
TTC GGG GTG GAG GAG GTC GTC CTC GTC CTG GAG GGA GAA AAA AAA AGC CTG AGC CCA AGG  1500
phe gly val glu glu val val leu val leu glu gly glu lys lys ser leu ser pro arg  (457)
```

FIG.4B-1

```
CCC CGC CCG GCC CCA CCT CCT GAA GCG CCC GCA CCC CCG GGC CCT CCC GAG GAG GAG GTA  1560
pro arg pro ala pro pro pro glu ala pro ala pro pro gly pro pro glu glu glu val  (477)

GAG GCG GAG GAA GCG GCG GAG GAG GCC CCG GAG GAG GCC TTG AGG CGG GTG GTC CGC CTC  1620
glu ala glu glu ala ala glu glu ala pro glu glu ala leu arg arg val val arg leu  (497)

CTG GGG GGG CGG GTG CTC TGG GTG CGG CGG CCC AGG ACC CGG GAG GCG CCG GAG GAG GAA  1680
leu gly gly arg val leu trp val arg arg pro arg thr arg glu ala pro glu glu glu  (517)

CCC CTG AGC CAA GAC GAG ATA GGG GGT ACT GGT ATA TAA    TGGGGGCATG   ACGCGGACCAC  1740
pro leu ser gln asp glu ile gly gly thr gly ile  *                               (529)

CGACCTCGGA   CAAGAGACCG   TGGACAACAT   CCTCAAGCGC   CTCCGCCGTA   TTGAGGGCCA  1820

GGTGCGGGGG   CTCCAGAAGA   TGGTGGCCGA   GGGCCGCCCC   TGCGACGAGG   TCCTCACCCA  1880

GATGACCGCC   ACCAAGAAGG   CCATGGAGGC   GGCGGCCACC   CTGATCCTCC   ACGAGTTCCT  1940

GAACGTCTGC   GCCGCCGAGG   TCTCCGAGGG   CAAGGTGAAC   CCCAAGAAGC   CCGAGGAGAT  2000

CGCCACCATG   CTGAAGAACT   TCATCTA                                           2027
```

FIG.4B-2

```
                GTG AGC GCC CTC TAC CGC CGC TTC CGC CCC CTC ACC TTC CAG GAG GTG GTG    51
GGG CAG GAG CAC GTG AAG GAG CCC CTC CTC AAG GCC ATC CGG GAG GGG AGG CTC GCC CAG   111
GCC TAC CTC TTC TCC GGG CCC AGG GGC GTG GGC AAG ACC ACC ACG GCG AGG CTC CTC GCC   171
ATG GCG GTG GGG TGC CAG GGG GAA GAC CCC CCT TGC GGG GTC TGC CCC CAC TGC CAG GCG   231
GtG CAG AGG GGC GCC CAC CCG GAC GTG GTG GAC ATT GAC GCC GCC AGC AAC AAC TCC GTG   291
GAG GAC GTG CGG GAG CTG AGG GAA AGG ATC CAC CTC GCC CCC CTC TCT GCC CCC AGG AAG   351
GTC TTC ATC CTG GAC GAG GCC CAC ATG CTC TCC AAA AGC GCC TTC AAC GCC CTC CTC AAG   411
ACC CTG GAG GAG CCC CCG CCC CAC GTC CTC TTC GTC TTC GCC ACC ACC GAG CCC GAG AGG   471
ATG CCC CCC ACC ATC CTC TCC CGC ACC CAG CAC TTC CGC TTC CGC CGC CTC ACG GAG GAG   531
GAG ATC GCC TTT AAG CTC CGG CGC ATC CTG GAG GCC GTG GGG CGG GAG GCG GAG GAG GAG   591
GCC CTC CTC CTC CTC GCC CGC CTG GCG GAC GGG GCC CTT AGG GAC GCG GAA AGC CTC CTG   651
GAG CGC TTC CTC CTC CTG GAA GGC CCC CTC ACC CGG AAG GAG GTG GAG CGC GCC CTA GGC   711
TCC CCC CCA GGG ACC GGG GTG GCC GAG ATC GCC GCC TCC CTC GCG AGG GGG AAA ACG GCG   771
GAG GCC CTG GGC CTC GCC CGG CGC CTC TAC GGG GAA GGG TAC GCC CCG AGG AGC CTG GTC   831
TCG GGC CTT TTG GAG GTG TTC CGG GAA GGC CTC TAC GCC GCC TTC GGC CTC GCG GGA ACC   891
CCC CTT CCC GCC CCG CCC CAG GCC CTG ATC GCC GCC ATG ACC GCC CTG GAC GAG GCC ATG   951
GAG CGC CTC GCC CGC CGC TCC GAC GCC TTA AGC CTG GAG GTG GCC CTC CTG GAG GCG GGA  1011
AGG GCC CTG GCC GCC GAG GCC CTA CCC CAG CCC ACG GGC GCT CCT TCC CCA GAG GTC GGC  1071
CCC AAG CCG GAA AGC CCC CCG ACC CCG GAA CCC CCA AGG CCC GAG GAG GCG CCC GAC CTG  1131
CGG GAG CGG TGG CGG GCC TTC CTC GAG GCC CTC AGG CCC ACC CTA CGG GCC TTC GTG CGG  1191
GAG GCC CGC CCG GAG GTC CGG GAA GGC CAG CTC TGC CTC GCT TTC CCC GAG GAC AAG GCC  1251
TTC CAC TAC CGC AAG GCC TCG GAA CAG AAG GTG AGG CTC CTC CCC CTG GCC CAG GCC CAT  1311
TTC GGG GTG GAG GAG GTC GTC CTC GTC CTG GAG GGA GAA AAA AAA AGC CTG AGC CCA AGG  1371
CCC CGC CCG GCC CCA CCT CCT GAA GCG CCC GCA CCC CCG GGC CCT CCC GAG GAG GAG GTA  1431
GAG GCG GAG GAA GCG GCG GAG GAG GCC CCG GAG GAG GCC TTG AGG CGG GTG GTC CGC CTC  1491
CTG GGG GGG CGG GTG CTC TGG GTG CGG CGG CCC AGG ACC CGG GAG GCG CCG GAG GAG GAA  1551
          CCC CTG AGC CAA GAC GAG ATA GGG GGT ACT GGT ATA TAA  (1590)
```

FIG.4C

Met ser ala leu tyr arg arg phe arg pro leu thr phe gln glu val val gly gln glu 20
his val lys glu pro leu leu lys ala ile arg glu gly arg leu ala gln ala tyr leu 40
phe ser gly pro arg gly val gly lys thr thr thr ala arg leu leu ala met ala val 60
gly cys gln gly glu asp pro pro cys gly val cys pro his cys gln ala val gln arg 80
gly ala his pro asp val val asp ile asp ala ala ser asn asn ser val glu asp val 100
arg glu leu arg glu arg ile his leu ala pro leu ser ala pro arg lys val phe ile 120
leu asp glu ala his met leu ser lys ser ala phe asn ala leu leu lys thr leu glu 140
glu pro pro pro his val leu phe val phe ala thr thr glu pro glu arg met pro pro 160
thr ile leu ser arg thr gln his phe arg phe arg arg leu thr glu glu glu ile ala 180
phe lys leu arg arg ile leu glu ala val gly arg glu ala glu glu glu ala leu leu 200
leu leu ala arg leu ala asp gly ala leu arg asp ala glu ser leu leu glu arg phe 220
leu leu leu glu gly pro leu thr arg lys glu val glu arg ala leu gly ser pro pro 240
gly thr gly val ala glu ile ala ala ser leu ala arg gly lys thr ala glu ala leu 260
gly leu ala arg arg leu tyr gly glu gly tyr ala pro arg ser leu val ser gly leu 280
leu glu val phe arg glu gly leu tyr ala ala phe gly leu ala gly thr pro leu pro 300
ala pro pro gln ala leu ile ala ala met thr ala leu asp glu ala met glu arg leu 320
ala arg arg ser asp ala leu ser leu glu val ala leu leu glu ala gly arg ala leu 340
ala ala glu ala leu pro gln pro thr gly ala pro ser pro glu val gly pro lys pro 360
glu ser pro pro thr pro glu pro pro arg pro glu glu ala pro asp leu arg glu arg 380
trp arg ala phe leu glu ala leu arg pro thr leu arg ala phe val arg glu ala arg 400
pro glu val arg glu gly gln leu cys leu ala phe pro glu asp lys ala phe his tyr 420
arg lys ala ser glu gln lys val arg leu leu pro leu ala gln ala his phe gly val 440
glu glu val val leu val leu glu gly glu lys lys ser leu ser pro arg pro arg pro 460
ala pro pro pro glu ala pro ala pro pro gly pro pro glu glu glu val glu ala glu 480
glu ala ala glu glu ala pro glu glu ala leu arg arg val val arg leu leu gly gly 500
arg val leu trp val arg arg pro arg thr arg glu ala pro glu glu glu pro leu ser 520
gln asp glu ile gly gly thr gly ile 529

FIG.4D

Met ser ala leu tyr arg arg phe arg pro leu thr phe gln glu val val gly gln glu 20
his val lys glu pro leu leu lys ala ile arg glu gly arg leu ala gln ala tyr leu 40
phe ser gly pro arg gly val gly lys thr thr thr ala arg leu leu ala met ala val 60
gly cys gln gly glu asp pro pro cys gly val cys pro his cys gln ala val gln arg 80
gly ala his pro asp val val asp ile asp ala ala ser asn asn ser val glu asp val 100
arg glu leu arg glu arg ile his leu ala pro leu ser ala pro arg lys val phe ile 120
leu asp glu ala his met leu ser lys ser ala phe asn ala leu leu lys thr leu glu 140
glu pro pro pro his val leu phe val phe ala thr thr glu pro glu arg met pro pro 160
thr ile leu ser arg thr gln his phe arg phe arg arg leu thr glu glu glu ile ala 180
phe lys leu arg arg ile leu glu ala val gly arg glu ala glu glu glu ala leu leu 200
leu leu ala arg leu ala asp gly ala leu arg asp ala glu ser leu leu glu arg phe 220
leu leu leu glu gly pro leu thr arg lys glu val glu arg ala leu gly ser pro pro 240
gly thr gly val ala glu ile ala ala ser leu ala arg gly lys thr ala glu ala leu 260
gly leu ala arg arg leu tyr gly glu gly tyr ala pro arg ser leu val ser gly leu 280
leu glu val phe arg glu gly leu tyr ala ala phe gly leu ala gly thr pro leu pro 300
ala pro pro gln ala leu ile ala ala met thr ala leu asp glu ala met glu arg leu 320
ala arg arg ser asp ala leu ser leu glu val ala leu leu glu ala gly arg ala leu 340
ala ala glu ala leu pro gln pro thr gly ala pro ser pro glu val gly pro lys pro 360
glu ser pro pro thr pro glu pro pro arg pro glu glu ala pro asp leu arg glu arg 380
trp arg ala phe leu glu ala leu arg pro thr leu arg ala phe val arg glu ala arg 400
pro glu val arg glu gly gln leu cys leu ala phe pro glu asp lys ala phe his tyr 420
arg lys ala ser glu gln lys val arg leu leu pro leu ala gln ala his phe gly val 440
glu glu val val leu val leu glu gly glu lys lys lys pro asp pro lys ala pro pro 460
gly pro thr ser 464

FIG.4E

Met ser ala leu tyr arg arg phe arg pro leu thr phe gln glu val val gly gln glu 20
his val lys glu pro leu leu lys ala ile arg glu gly arg leu ala gln ala tyr leu 40
phe ser gly pro arg gly val gly lys thr thr thr ala arg leu leu ala met ala val 60
gly cys gln gly glu asp pro pro cys gly val cys pro his cys gln ala val gln arg 80
gly ala his pro asp val val asp ile asp ala ala ser asn asn ser val glu asp val 100
arg glu leu arg glu arg ile his leu ala pro leu ser ala pro arg lys val phe ile 120
leu asp glu ala his met leu ser lys ser ala phe asn ala leu leu lys thr leu glu 140
glu pro pro pro his val leu phe val phe ala thr thr glu pro glu arg met pro pro 160
thr ile leu ser arg thr gln his phe arg phe arg arg leu thr glu glu glu ile ala 180
phe lys leu arg arg ile leu glu ala val gly arg glu ala glu glu glu ala leu leu 200
leu leu ala arg leu ala asp gly ala leu arg asp ala glu ser leu leu glu arg phe 220
leu leu leu glu gly pro leu thr arg lys glu val glu arg ala leu gly ser pro pro 240
gly thr gly val ala glu ile ala ala ser leu ala arg gly lys thr ala glu ala leu 260
gly leu ala arg arg leu tyr gly glu gly tyr ala pro arg ser leu val ser gly leu 280
leu glu val phe arg glu gly leu tyr ala ala phe gly leu ala gly thr pro leu pro 300
ala pro pro gln ala leu ile ala ala met thr ala leu asp glu ala met glu arg leu 320
ala arg arg ser asp ala leu ser leu glu val ala leu leu glu ala gly arg ala leu 340
ala ala glu ala leu pro gln pro thr gly ala pro ser pro glu val gly pro lys pro 360
glu ser pro pro thr pro glu pro pro arg pro glu glu ala pro asp leu arg glu arg 380
trp arg ala phe leu glu ala leu arg pro thr leu arg ala phe val arg glu ala arg 400
pro glu val arg glu gly gln leu cys leu ala phe pro glu asp lys ala phe his tyr 420
arg lys ala ser glu gln lys val arg leu leu pro leu ala gln ala his phe gly val 440
glu glu val val leu val leu glu gly glu lys lys lys ala 454

FIG.4F

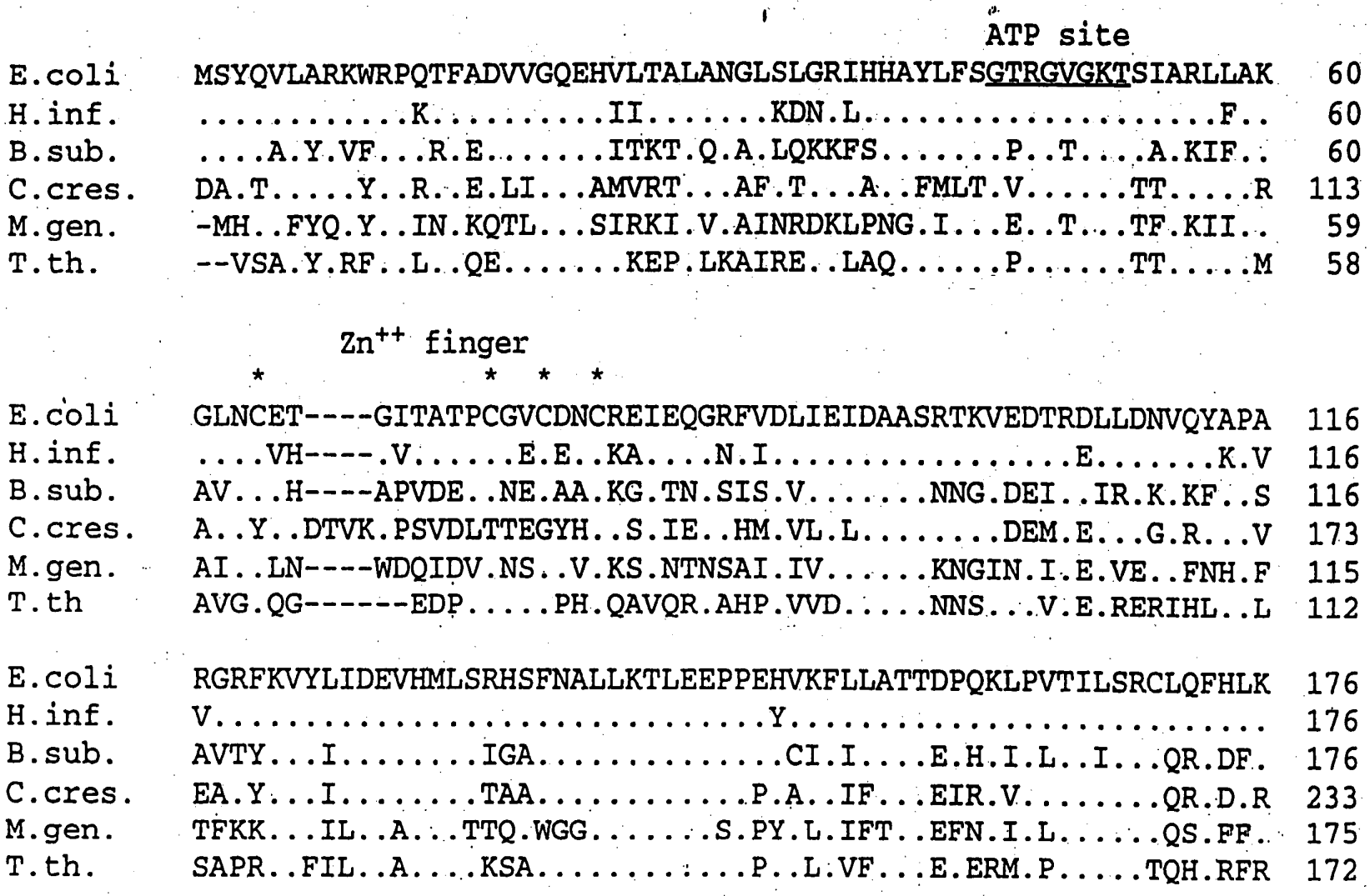

ATP site
E.coli MSYQVLARKWRPQTFADVVGQEHVLTALANGLSLGRIHHAYLFSGTRGVGKTSIARLLAK 60
H.inf. .....K.....II.....KDN.L.....F.. 60
B.sub. ....A.Y.VF...R.E.....ITKT.Q.A.LQKKFS.....P..T....A.KIF.. 60
C.cres. DA.T.....Y..R..E.LI...AMVRT...AF.T...A..FMLT.V.....TT.....R 113
M.gen. -MH..FYQ.Y..IN.KQTL...SIRKI.V.AINRDKLPNG.I...E..T...TF.KII.. 59
T.th. --VSA.Y.RF..L..QE.....KEP.LKAIRE..LAQ.....P.....TT.....M 58
Zn++ finger
E.coli GLNCET----GITATPCGVCDNCREIEQGRFVDLIEIDAASRTKVEDTRDLLDNVQYAPA 116
H.inf. ....VH----.V.....E.E..KA....N.I.....E.....K.V 116
B.sub. AV...H----APVDE..NE.AA.KG.TN.SIS.V.....NNG.DEI..IR.K.KF..S 116
C.cres. A..Y..DTVK.PSVDLTTEGYH..S.IE..HM.VL.L.....DEM.E...G.R...V 173
M.gen. AI..LN----WDQIDV.NS..V.KS.NTNSAI.IV.....KNGIN.I.E.VE..FNH.F 115
T.th AVG.QG------EDP.....PH.QAVQR.AHP.VVD.....NNS...V.E.RERIHL..L 112
E.coli RGRFKVYLIDEVHMLSRHSFNALLKTLEEPPEHVKFLLATTDPQKLPVTILSRCLQFHLK 176
H.inf. V.....Y..... 176
B.sub. AVTY...I.....IGA.....CI.I....E.H.I.L..I...QR.DF. 176
C.cres. EA.Y...I.....TAA.....P.A..IF...EIR.V.....QR.D.R 233
M.gen. TFKK...IL..A...TTQ.WGG.....S.PY.L.IFT..EFN.I.L.....QS.FF. 175
T.th. SAPR..FIL..A....KSA.....P..L.VF...E.ERM.P.....TQH.RFR 172

FIG.5A

```
E.coli     ALDVEQIRHQLEHILNEEHIAHEPRALQLLARAAEGSLRDALSLTDQAIASGDGQ--VST  234
H.inf.     ...ET..SQH.A...TQ.N.PF.DP..VK..K..Q..I..S.........M..R.--.TN  234
B.sub.     RITSQA.VGRMNK.VDA.QLQV.EGS.EII.S..H.GM......L....SFSGDI--LKV  234
C.cres.    RVEPDVLVKHFDR.SAK.GARI.MD..A.I.......V..G...L....VQTERGQT.TS  293
M.gen.     KITSDL.LER.ND.AKK.K.KI.KD..IKI.DLSQ.....G...L..LAI.LIVKKL.LL  235
T.th.      R.TE.E.AFK.RR..EAVGREA.EE..L....L.D.A....E..LERFLLLEGP---LTR  229

E.coli     QAVSAMLGTLDDDQALSLVEAMVEANGERVMALINEAAARGIEWEALLVEMLGLLHRIAM  294
H.inf.     NV..N...L...NYSVDILY.LHQG...LL.RTLQRV.DAAGD.DK..G.CAEK..Q..L  294
B.sub.     EDALLIT.AVSQLYIGK.AKSLHDK.VSDALETL..LLQQ.KDPAK.IED.IFYFRDMLL  294
C.cres.    TV.RD...LA.RS.TIA.Y.HVMAGKTKDALEGFRALWGF.ADPAVVMLDV.DHC.AS.V  353
M.gen.     MLKKHLISLIEMQNL.L.KQFYQ.I                                      260
T.th.      KE.ERA..SPPGTGVAEIAASLARGKTAEALG.ARRLYGE.YAPRS.VSGL.EVFREGLY  289
```

FIG.5B

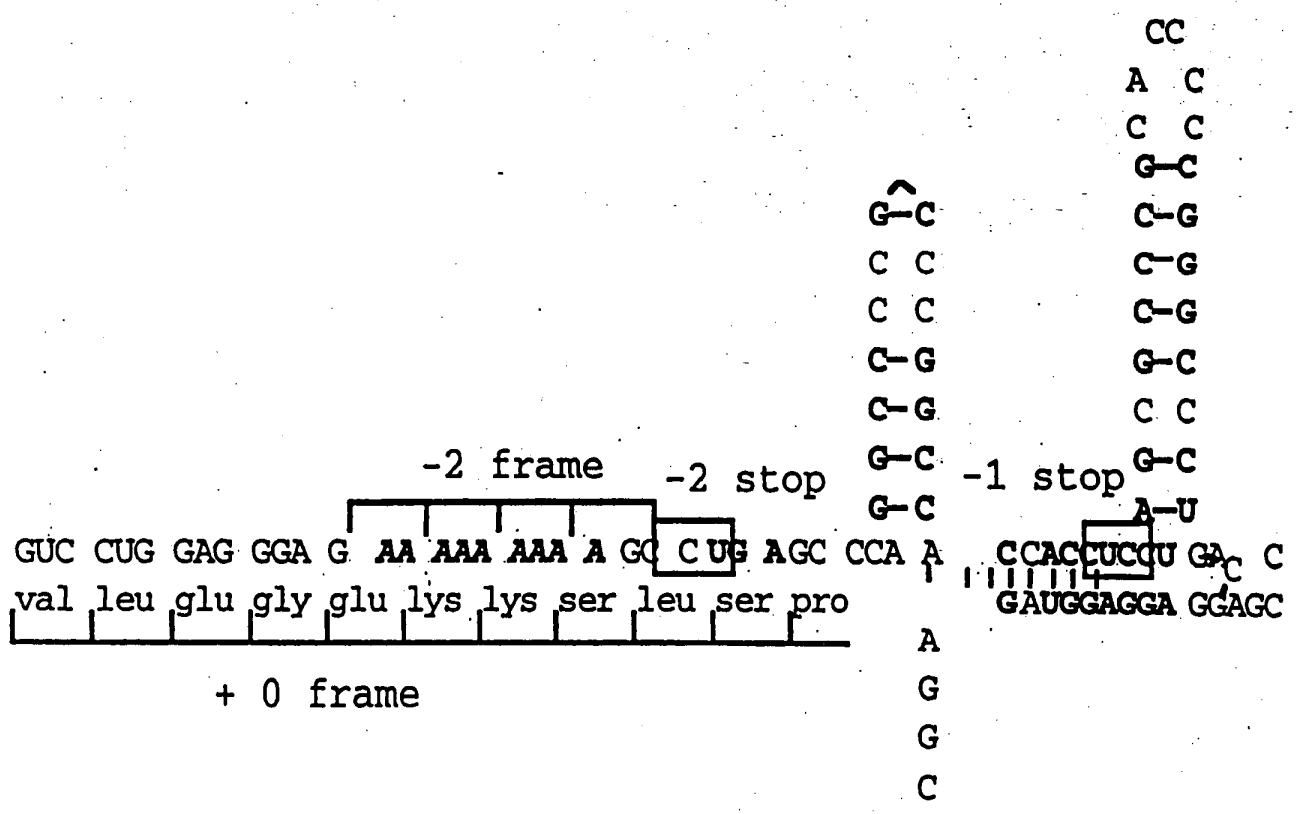

-2 frame
-2 stop
-1 stop
GUC CUG GAG GGA G AA AAA AAA A GC CUG AGC CCA A CCACCUCCU GA C C
val leu glu gly glu lys lys ser leu ser pro
GAUGGAGGA GGAGC
+ 0 frame

FIG.6

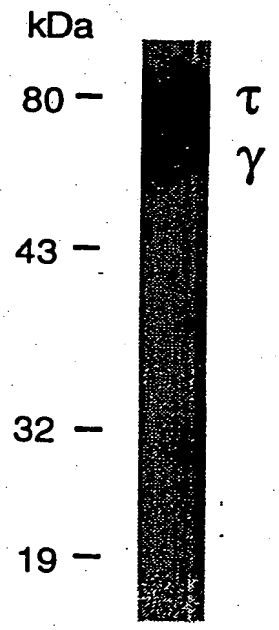
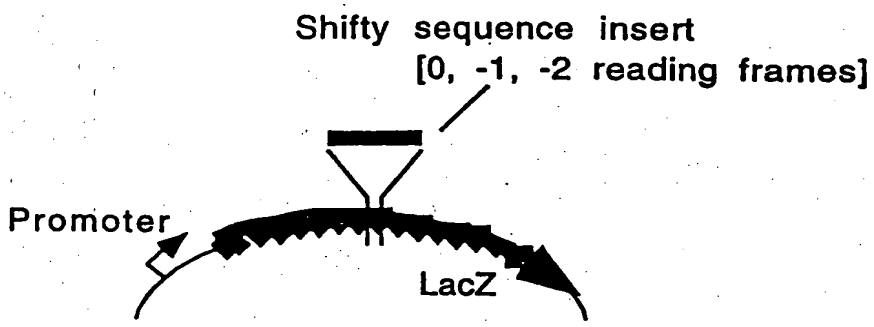


FIG.7

FIG.8A



	Reading frame	Blue	White
Shifty sequence	0	+	
	- 1	+	
	- 2	+	
Mutant sequence	0	++	
	- 1		+
	- 2		+

FIG.8B

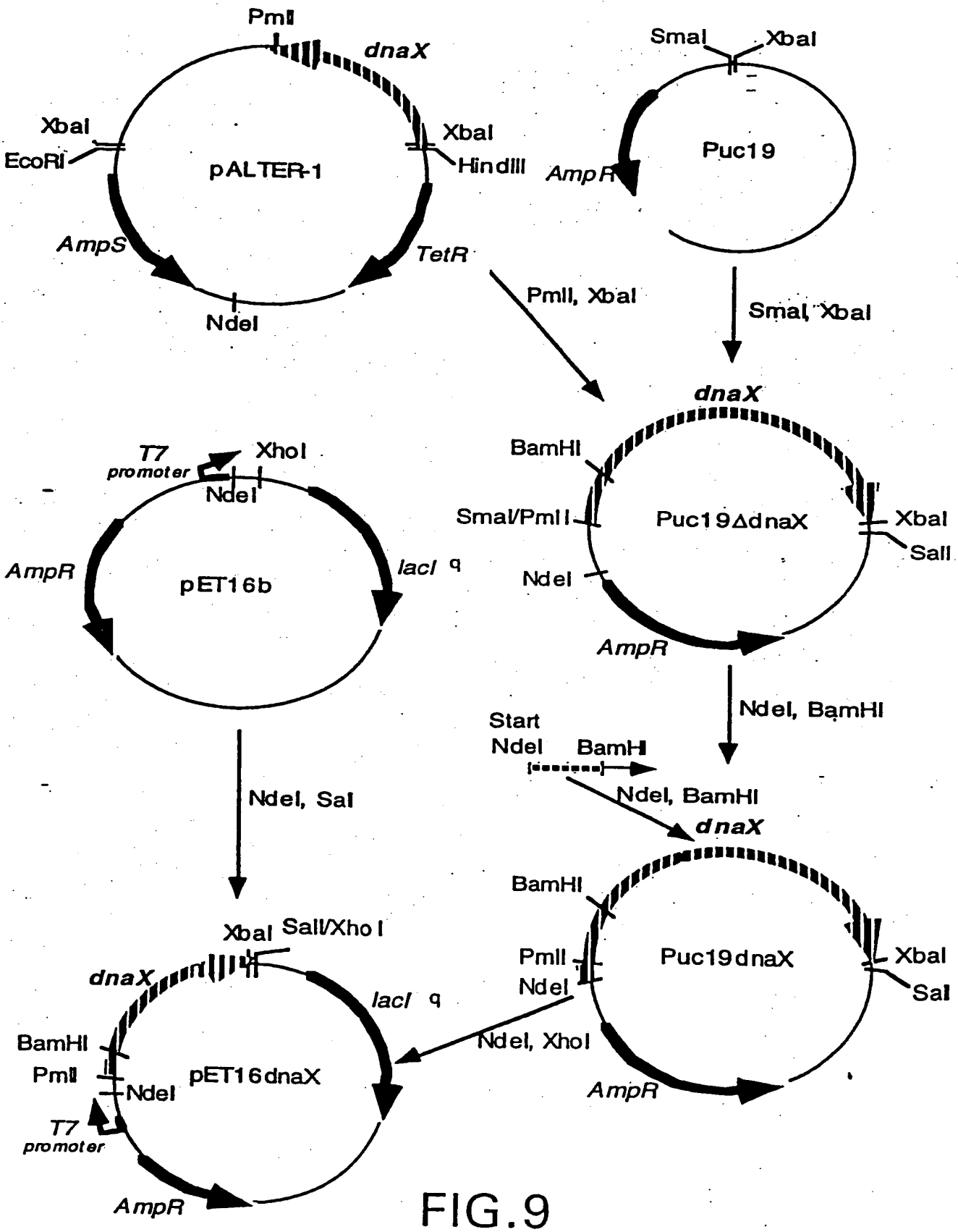

PmlI
dnaX
XbaI
EcoRI
pALTER-1
XbaI
HindIII
AmpS
TetR
NdeI
SmaI
XbaI
Puc19
AmpR
PmlI, XbaI
SmaI, XbaI
dnaX
BamHI
SmaI/PmlI
Puc19ΔdnaX
XbaI
SalI
NdeI
AmpR
T7
promoter
XhoI
NdeI
AmpR
pET16b
lacI q
NdeI, BamHI
Start
NdeI
BamHI
NdeI, BamHI
dnaX
BamHI
PmlI
NdeI
Puc19dnaX
XbaI
Sal
NdeI, XhoI
AmpR
NdeI, SalI
XbaI
SalI/XhoI
dnaX
lacI q
BamHI
PmlI
NdeI
pET16dnaX
T7
promoter
AmpR

FIG.9

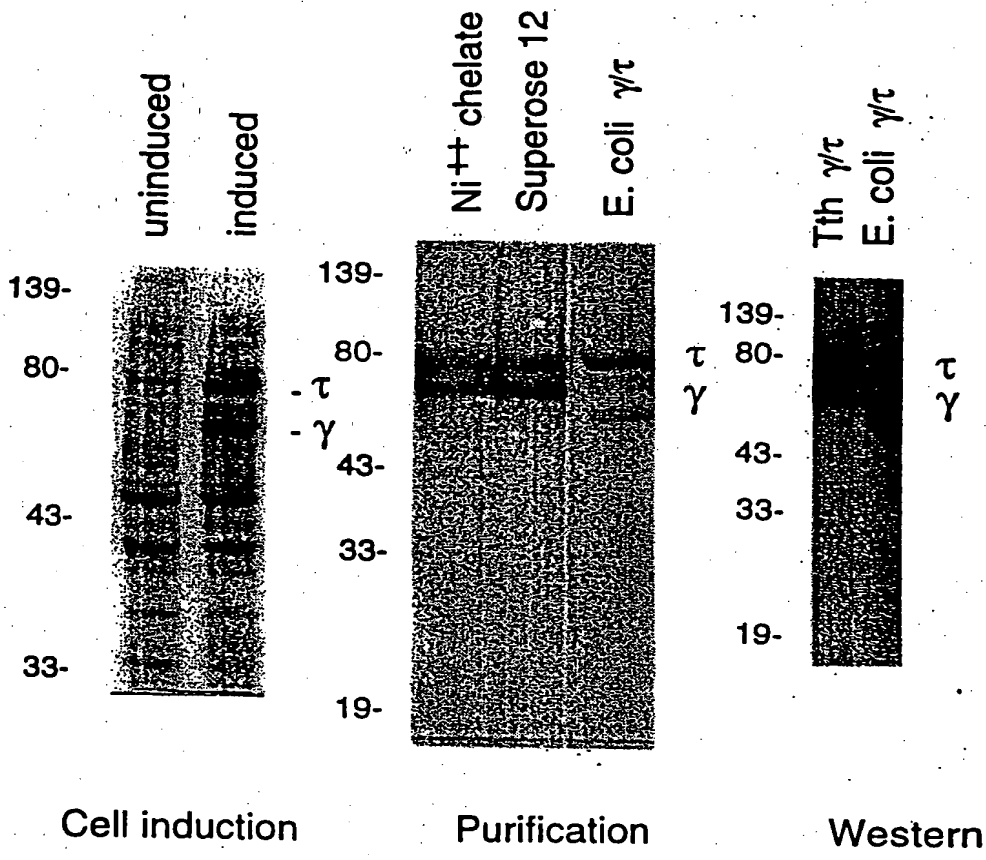


FIG.10A FIG.10B FIG.10C

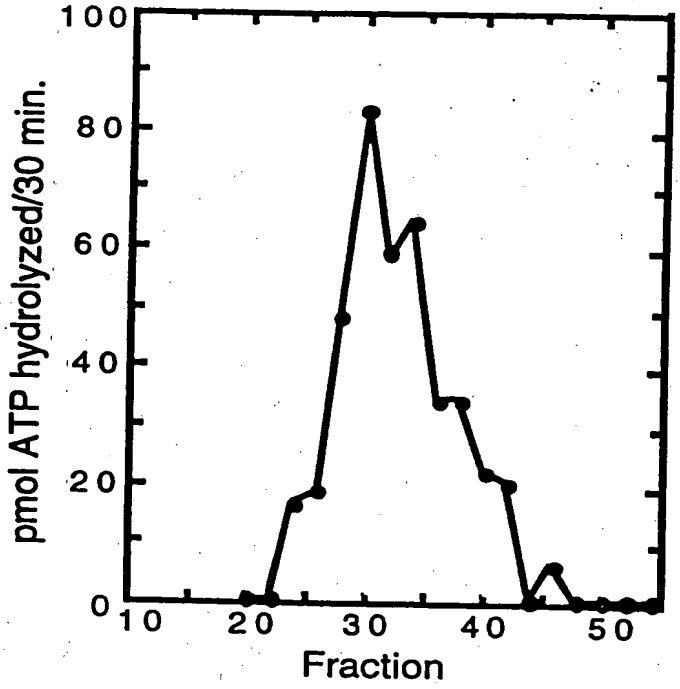


FIG.11A

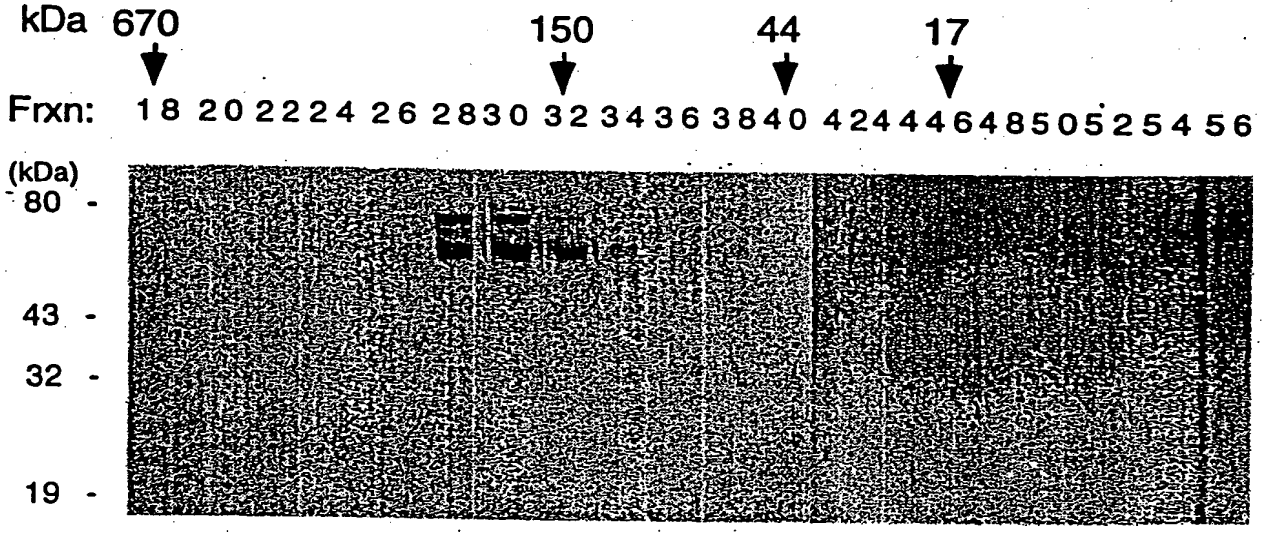


FIG.11B

FIG.12A

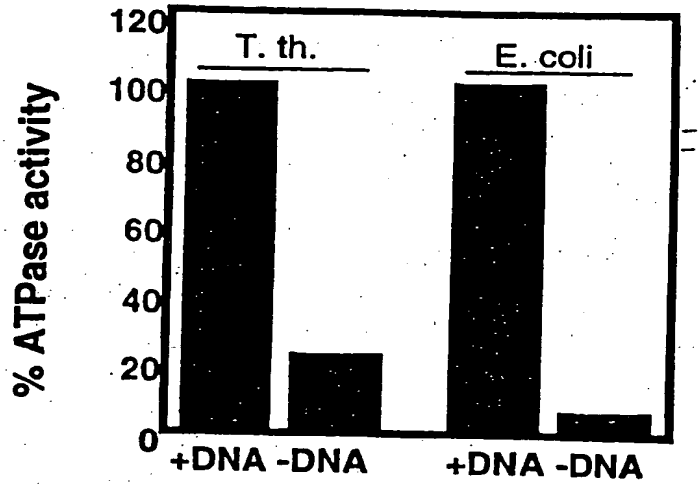


FIG.12B

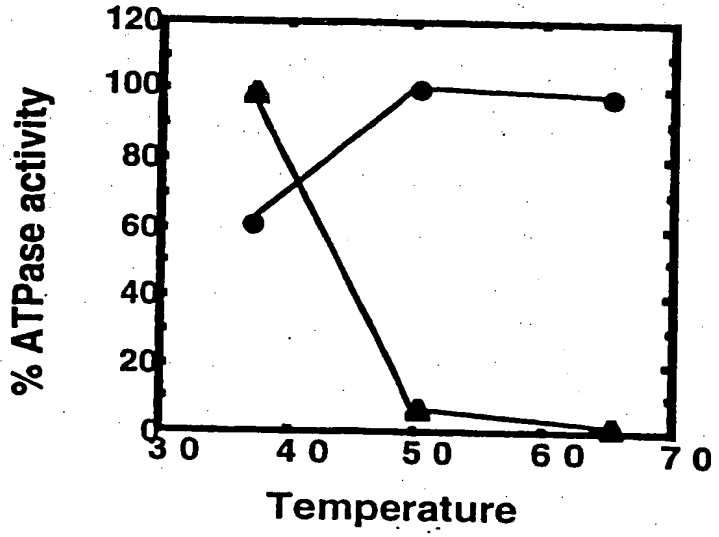


FIG.12C

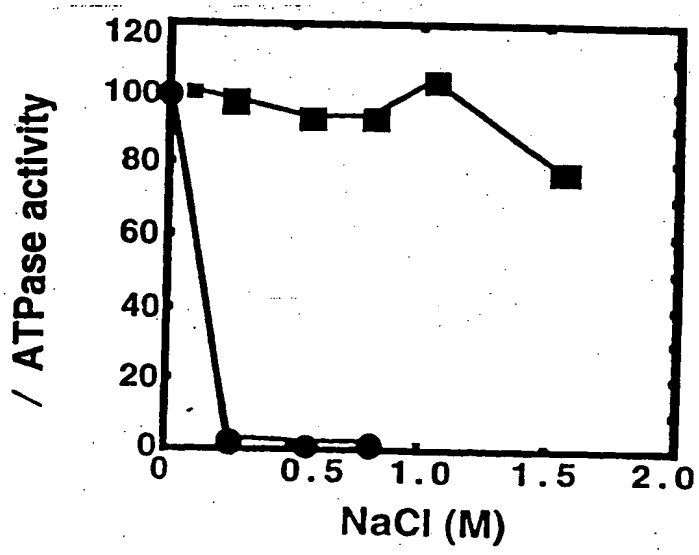


FIG.13A

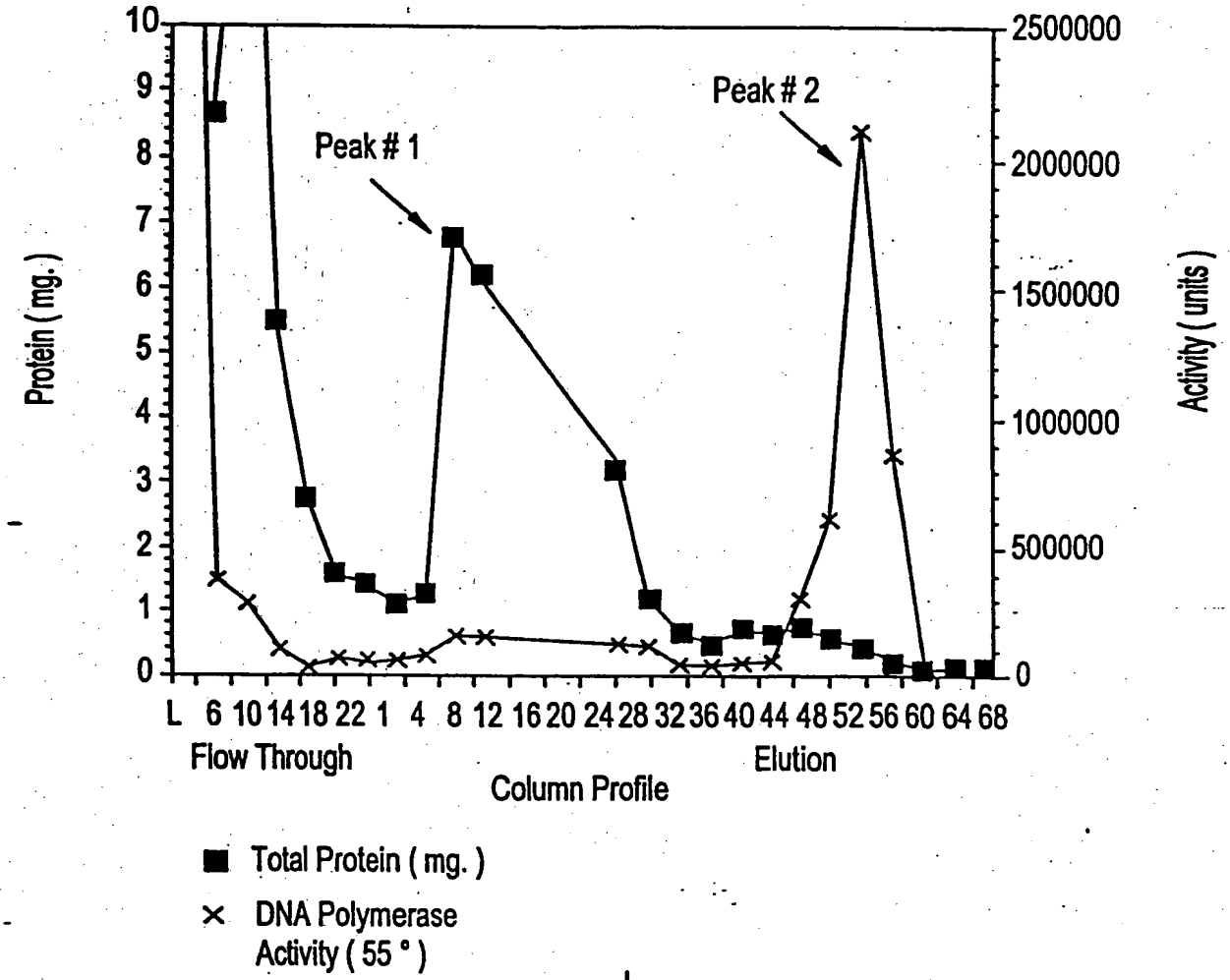


FIG.13B

ATP Agarose Step Column

FIG.13C

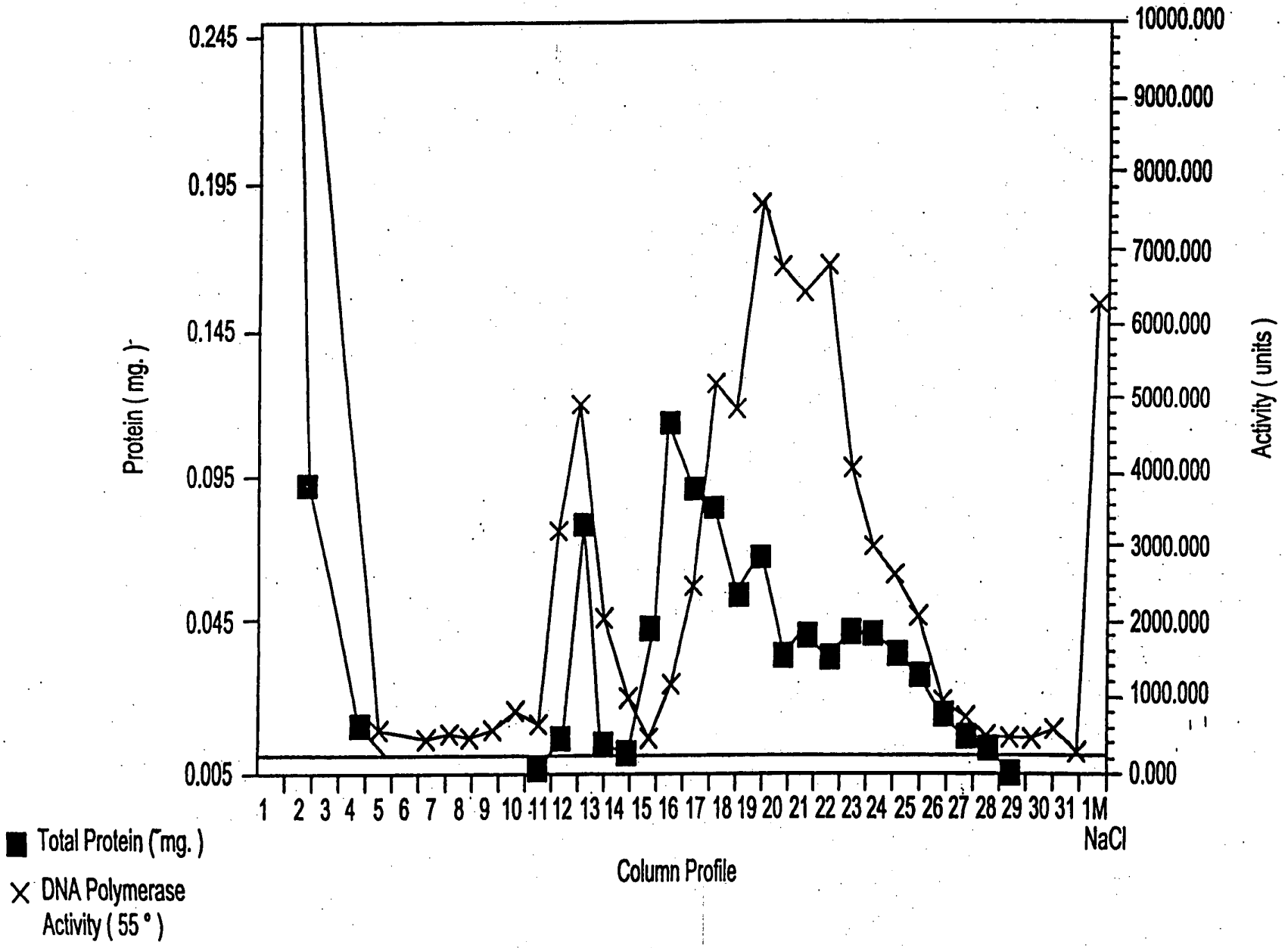

0.245
0.195
0.145
0.095
0.045
0.005
Protein (mg.)
10000.000
9000.000
8000.000
7000.000
6000.000
5000.000
4000.000
3000.000
2000.000
1000.000
0.000
Activity (units)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 1M
NaCl
Column Profile
Total Protein (mg.)
DNA Polymerase Activity (55 °)

FIG. 14A

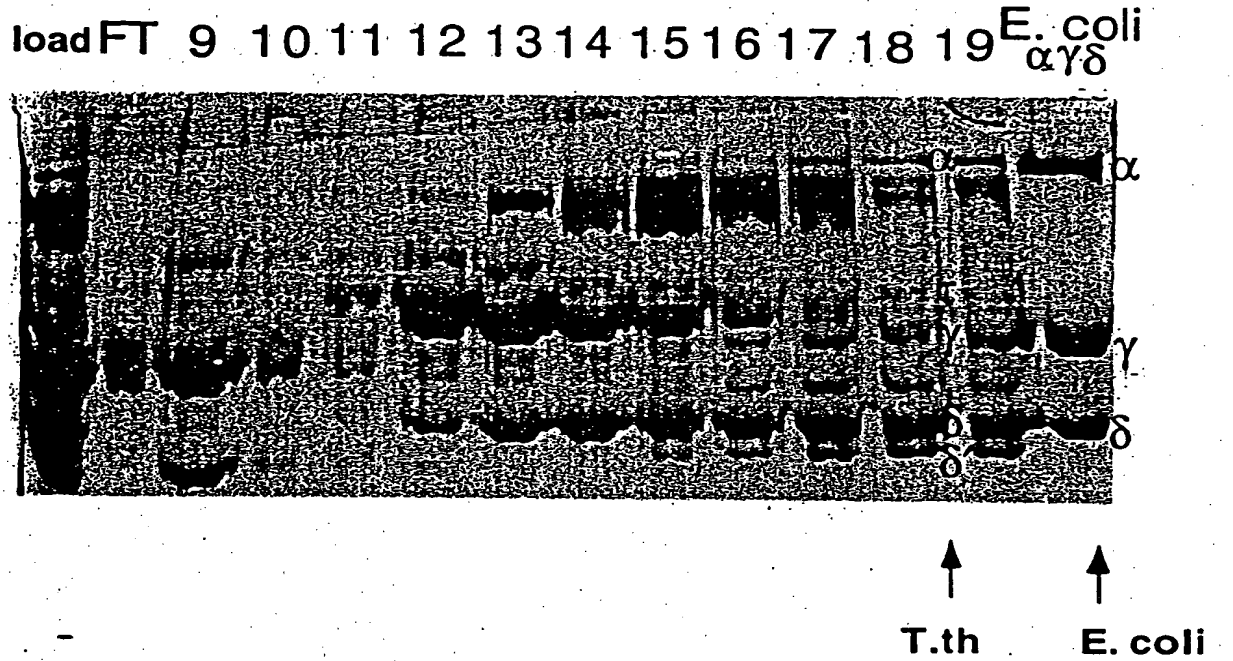
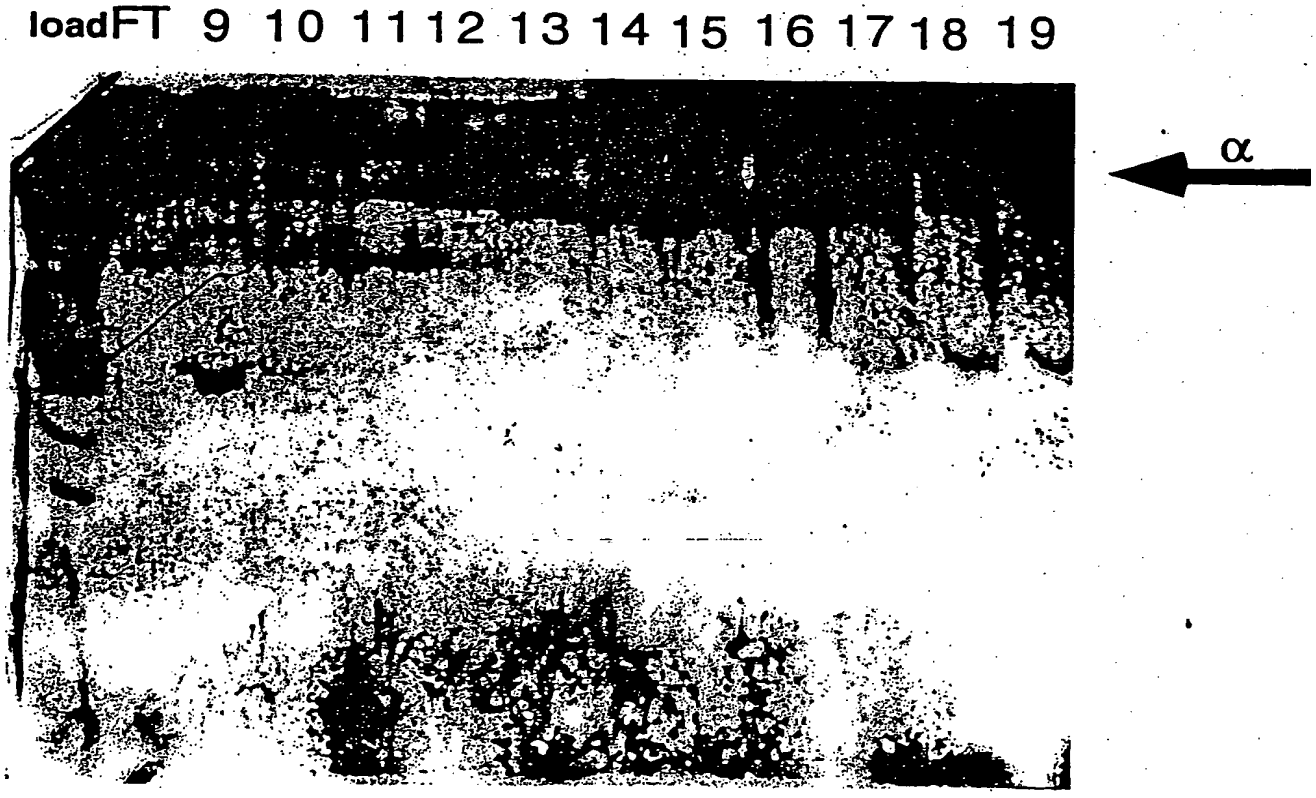


FIG. 14B



Alignment of TTH1 with alphas subunits of other organisms.

E.coli	DRYFLELIRTGRPDEESYLHAAVELAEARGLPVV	197	(ID#72)
V.chol.	DHFYLELIRTGRADEESYLHFALDVAEQYDLPVV	197	(ID#73)
H.inf.	DHFYLALSRTGRPNEERYIQAALKLAERCDLPLV	197	(ID#74)
R.prow.	DRFYFEIMRHDLPEEQFIENSYIQIASELSIPIV	195	(ID#75)
H.pyl.	DDFYLEIMRHGILDQRFIDEQVIKMSLETGLKII	213	(ID#76)
S.sp.	DDYYLEIQDHGSVEDRLVNINLVKIAQELDIKIV	202	(ID#77)
M.tub.	DNYFLELMDHGLTIERRVRDGLLEIGRALNIPPL	220	(ID#78)
T.th.	FFIEIQNHGLSEQK		(ID#61)

FIG.15A

Alignment of TTH2 with alphas subunits of other organisms.

E.coli	NKRRAKNGEPPLDIAAIPLDDKKSFDMLQRSETTAVFQLESRGMKD	618	(ID#79)
V.chol.	NPRLKKAGKPPVRIEAIPLDDARSFRNLQDAKTTAVFQLESRGMKE	618	(ID#80)
H.inf.	NVRMVREGKPRVDIAAIPLDDPESFELLKRSETTAVFQLESRGMKD	618	(ID#81)
R.prow.	CKKLLKEQGIKIDFDDMTFDDKKTYQMLCKGKGVGVFQFESIGMKD	624	(ID#82)
H.pyl.	LKIIKTQHKISVDFLSLDMDDPKVYKTIQSGDTVGIFQIES-GMFQ	648	(ID#83)
S.sp.	QERKALQIRARTGSKKLPDDVKKTHKLLEAGDLEGIFQLESQGMKQ	643	(ID#84)
M.tub.	IDNVRANRGIDLDLESVPLDDKATYELLGRGDTLGVFQLDGGPMRD	646	(ID#85)
T.th.	RVELDYDALTLDD		(ID#60)

FIG.15B

ATGGGCCGGGAGCTCCGCTTCGCCCACCTCCACCAGCACA
CCCAGTTCTCCCTCCTGGACGGGGCGGCGAAGCTTTCCGA
CCTCCTCAAGTGGGTCAAGGAGACGACCCCCGAGGACCCC 120
GCCTTGGCCATGACCGACCACGGCAACCTCTTCGGGGCCG
TGGAGTTCTACAAGAAGGCCACCGAAATGGGCATCAAGCC
CATCCTGGGCTACGAGGCCTACGTGGCGGCGGAAAGCCGC 240
TTTGACCGCAAGCGGGGAAAGGGCCTAGACGGGGGCTACT
TTCACCTCACCCTCCTCGCCAAGGACTTCACGGGGTACCA
GAACCTGGTGCGCCTGGCGAGCCGGGCTTACCTGGAGGGG 360
TTTTACGAAAAGCCCCGGATTGACCGGGAGATCCTGCGCG
AGCACGCCGAGGGCCTCATCGCCCTCTCGGGGTGCCTCGG
GGCGGAGATCCCCCAGTTCATCCTCCAGGACCGTCTGGAC 480
CTGGCCGAGGCCCGGCTCAACGAGTACCTCTCCATCTTCA
AGGACCGCTTCTTCATCGAGATCCAGAACCACGGCCTCCC
CGAGCAGAAAAAGGTCAACGAGGTCCTCAAGGAGTTCGCC 600
CGAAAGTACGGCCTGGGGATGGTGGCCACCAACGACGGCC
ATTACGTGAGGAAGGAGGACGCCCGCGCCCACGAGGTCCT
CCTCGCCATCCAGTCCAAGAGCACCCTGGACGACCCCGGG 720
CGCTGGCGCTTCCCCTGCGACGAGTTCTACGTGAAGACCC
CCGAGGAGATGCGGGCCATGTTCCCCGAGGAGGAGTGGGG
GGACGAGCCCTTTGACAACACCGTGGAGATCGCCCGCATG 840
TGCAACGTGGAGCTGCCCATCGGGGACAAGATGGTCTACC
GAATCCCCCGCTTCCCCCTCCCCGAGGGGCGGACCGAGGC
CCAGTACCTCATGGAGCTCACCTTCAAGGGGCTCCTCCGC 960
CGCTACCCGGACCGGATCACCGAGGGCTTCTACCGGGAGG
TCTTCCGCCTTTTGGGGAAGCTTCCCCCCCACGGGGACGG
GGAGGCCTTGGCCGAGGCCTTGGCCCAGGTGGAGCGGGAG 1080
GCTTGGGAGAGGCTCATGAAGAGCCTCCCCCCTTTGGCCG
GGGTCAAGGAGTGGACGGCGGAGGCCATTTTCCACCGGGC
CCTTTACGAGCTTTCCGTGATAGAGCGCATGGGGTTTCCC 1200
GGCTACTTCCTCATCGTCCAGGACTACATCAACTGGGCCC
GGAGAAACGGCGTCTCCGTGGGGCCCGGCAGGGGGAGCGC
CGCCGGGAGCCTGGTGGCCTACGCCGTGGGGATCACCAAC 1320
ATTGACCCCCTCCGCTTCGGCCTCCTCTTTGAGCGCTTCC
TGAACCCGGAGAGGGTCTCCATGCCCGACATTGACACGGA
CTTCTCCGACCGGGAGCGGGACCGGGTGATCCAGTACGTG 1440
CGGGAGCGCTACGGCGAGGACAAGGTGGCCCAGATCGGCA
CCCTGGGAAGCCTCGCCTCCAAGGCCGCCCTCAAGGACGT
GGCCCGGGTCTACGGCATCCCCCACAAGAAGGCGGAGGAA 1560
TTGGCCAAGCTCATCCCGGTGCAGTTCGGGAAGCCCAAGC
CCCTGCAGGAGGCCATCCAGGTGGTGCCGGAGCTTAGGGC
GGAGATGGAGAAGGACCCCAAGGTGCGGGAGGTCCTCGAG 1680
GTGGCCATGCGCCTGGAGGGCCTGAACCGCCACGCCTCCG
TCCACGCCGCCGGGGTGGTGATCGCCGCCGAGCCCCTCAC
GGACCTCGTCCCCCTCATGCGCGACCAGGAAGGGCGGCCC 1800
GTCACCCAGTACGACATGGGGGCGGTGGAGGCCTTGGGGC
TTTTGAAGATGGACTTTTTGGGCCTCCGCACCCTCACCTT

FIG. 16A

CCTGGACGAGGTCAAGCGCATCGTCAAGGCGTCCCAGGGG 1920
GTGGAGCTGGACTACGATGCCCTCCCCCTGGACGACCCCA
AGACCTTCGCCCTCCTCTCCCGGGGGGAGACCAAGGGGGT
CTTCCAGCTGGAGTCGGGGGGGATGACCGCCACGCTCCGC 2040
GGCCTCAAGCCGCGGCGCTTTGAGGACCTGATCGCCATCC
TCTCCCTCTACCGCCCCGGGCCCATGGAGCACATCCCCAC
CTACATCCGCCGCCACCACGGGCTGGAGCCCGTGAGCTAC 2160
AGCGAGTTTCCCCACGCCGAGAAGTACCTAAAGCCCATCC
TGGACGAGACCTACGGCATCCCCGTCTACCAGGAGCAGAT
CATGCAGATCGCCTCGGCCGTGGCGGGGTACTCCCTGGGC 2280
GAGGCGGACCTCCTGCGGCGGTCCATGGGCAAGAAGAAGG
TGGAGGAGATGAAGTCCCACCGGGAGCGCTTCGTCCAGGG
GGCCAAGGAAAGGGGCGTGCCCGAGGAGGAGGCCAACCGC 2400
CTCTTTGACATGCTGGAGGCCTTCGCCAACTACGGCTTCA
ACAAATCCCACGCTGCCGCCTACAGCCTCCTCTCCTACCA
GACCGCCTACGTGAAGGCCCACTACCCCGTGGAGTTCATG 2520
GCCGCCCTCCTCTCCGTGGAGCGGCACGACTCCGACAAGG
TGGCCGAGTACATCCGCGACGCCCGGGCCATGGGCATAGA
GGTCCTTCCCCCGGACGTCAACCGCTCCGGGTTTGACTTC 2640
CTGGTCCAGGGCCGGCAGATCCTTTTCGGCCTCTCCGCGG
TGAAGAACGTGGGCGAGGCGGCGGCGGAGGCCATTCTCCG
GGAGCGGGAGCGGGGCGGCCCCTACCGGAGCCTCGGCGAC 2760
TTCCTCAAGCGGCTGGACGAGAAGGTGCTCAACAAGCGGA
CCCTGGAGTCCCTCATCAAGGCGGGCGCCCTGGACGGCTT
CGGGGAAAGGGCGCGGCTCCTCGCCTCCCTGGAAGGGCTC 2880
CTCAAGTGGGCGGCCGAGAACCGGGAGAAGGCCCGCTCGG
GCATGATGGGCCTCTTCAGCGAAGTGGAGGAGCCGCCTTT
GGCCGAGGCCGCCCCCCTGGACGAGATCACCCGGCTCCGC 3000
TACGAGAAGGAGGCCCTGGGGATCTACGTCTCCGGCCACC
CCATCTTGCGGTACCCCGGGCTCCGGGAGACGGCCACCTG
CACCCTGGAGGAGCTTCCCCACCTGGCCCGGGACCTGCCG 3120
CCCCGGTCTAGGGTCCTCCTTGCCGGGATGGTGGAGGAGG
TGGTGCGCAAGCCCACAAAGAGCGGCGGGATGATGGCCCG
CTTCGTCCTCTCCGACGAGACGGGGGCGCTTGAGGCGGTG 3240
GCATTCGGCCGGGCCTACGACCAGGTCTCCCCGAGGCTCA
AGGAGGACACCCCCGTGCTCGTCCTCGCCGAGGTGGAGCG
GGAGGAGGGGGGCGTGCGGGTGCTGGCCCAGGCCGTTTGG 3360
ACCTACGAGGAGCTGGAGCAGGTCCCCCGGGCCCTCGAGG
TGGAGGTGGAGGCCTCCCTCCTGGACGACCGGGGGGTGGC
CCACCTGAAAAGCCTCCTGGACGAGCACGCGGGGACCCTC 3480
CCCCTGTACGTCCGGGTCCAGGGCGCCTTCGGCGAGGCCC
TCCTCGCCCTGAGGGAGGTGCGGGTGGGGGAGGAGGCTGT
AGGCGGCCGCGTGGTTCCGGGCCTACCTCCTGCCCGACCG 3600
GGAGGTCCTTCTCCAGGGCGGCCAGGCGGGGGAGGCCCAG
GAGGCGGTGCCCTTCTAGGGGGTGGGCCGTGAGACCTAGC
GCCATCGTTCTCGCCGGGGGCAAGGAGGCCTGGGCCCGAC 3720
CCCTTTTGG

FIG. 16B

MGRELRFAHLHQHTQFSLLDGAPKLSDLLKWVEETTPEDP
ALAMTDHGNLFGAVEFYKKATEMGIKPILGYEAYVAAESR
FDRKRGKGLDGGYFHLTLLAKDFTGYQNLVRLASRAYLEG 120
FYEKPRIDREILREHAEGLIALSGCLGAEIPQFILQDRLD
LAEARLNEYLSIFKDRFFIEIQNHGLPEQKKVNEVLKEFA
RKYGLGMVATNDGHYVRKEDARAHEVLLAIQSKSTLDDPG 240
ALALPCEEFYVKTPEEMRAMFPEEEVGGRSPLTTPWRSPH
VQRGAAIGTRWSTRIPRFPLPEGRTEAQYLMELTFKGLLR
RYPDRITEGFYREVFRLSGKLPPHGDGEALAEALAQVERE 360
AWERLMKSLPPLAGVKEWTAEAIFHRALYELSAIERMGFP
GLLPHRPGLHQLGPEKGVSVGPGRGGAAGSLVAYAVGITN
IDPLRFGLLFERFLNPERVSMPDIDTDFSDRERDRVIQYV 480
RERYGEDKVAQIGTLGSLASKAALKEVARVYGIPRKKAEE
LAKLIPVQFGKPKPLQEAIQVVPELRAEMEKDPKVREVLE
VAMRLEGLNRHASVHAGRGGVFSEPLTDLVPLCATRKGGP 600
YTQYDMGAVEALGLLKMDFLGLRTLTFLDEVKRIVKASQG
VELDYDALPLDDPKTFALLSRGETKGVFQLESGGMTATLR
GLKPRRFEDLIAILSLYRPGPMEHIPTYIRRHHGLEPVSY 720
SEFPHAEKYLKPILDETYGIPVYQEQIMQIASAVAGYSLG
EADLLRRSMGKKKVEEMKSHRERFVQGAKERGVPEEEANR
LFDMLEAFANYGFNKSHAAAYSLLSYQTAYVKAHYPVEFM 840
AALLSVERHDSDKVAEYIRDARAMGIEVLPPDVNRSGFDF
LVQGRQILFGLSAVKNVGEAAAEAILRERERGGPYRSLGD
FLKRLDEKVLNKRTLESLIKAGALDGFGERARLLASLEGL 960
LKWAAENREKARSGMMGLFSEVEEPPLAEAAPLDEITRLR
YEKEALGIYVSGHPILRYPGLRETATCTLEELPHLARDLP
PRSRVLLAGMVEEVVRKPTKSGGMMARFVLSDETGALEAV 1080
AFGRAYDQVSPRLKEDTPVLVLAEVEREEGGVRVLAQAVW
TYQELEQVPRALEVEVEASLPDDRGVAHLKSLLDEHAGTL
PLYVRVQGAFGEALLALREVRVGEEALGALEAAGFPAYLL 1200
PNREVSPRLTGSGGPRGRALSTGLALKTYPIALPGGNEAL
ARPLL

FIG. 16C

```
          Start1                Start2                3'-Exo I
T.th.     VERVVRTLLDGRFLLEEGVGLWEWRYPFPLEGEAVVVLDLETTGLAG------LDEVIEVGLLRLEGG---RRLPF
D.rad.                                                PWPQDVVVFDLETTGFSPA-----SAAIVEIGAVRIVGGQIDETLKF
Bac.sub.  HGIKMIYGMEANLVDDGVPIAYNAAHRLLEEETYVVFDVETTGLSAV-----YDTIIELAAVKVKGGE--IIDKF
H.inf.                                              MINPNRQIVLDTETTGMNQLGAHYEGHCIIEIGAVELINRR-YTGNNX
E.c.                                               MSTAITRQIVLDTETTGMNQIGAHSEGHKIIEIGAVEVVNRR-LTGNNF
H.pyl.    NLEYLKACGLNFIETSENLITLKNLKTPLKDEVFSFIDLETTGSCPI-----KHEILEIGAVQVKGGE--IINRF

                                                                                    3'-Exo II
T.th.     QSLVR-PLPP---AEARSWNLT---GIPREALEEAPSLEEVLEKAYPLRGDATLVIHNAAFDLGFL-RPALEGLG
D.rad.    ETLVR-PTRPDGSMLSIPWQAQRVHGISDEMVRRAPAXKDVLPDFFDFVDGSAVVAHNVSFDGGFM-RAGAERLG
Bac.sub.  EAFAN-PHRP---LSATIIELT---GITDDMLQDAPDVVDVIRDFREWIGDDILVAHNASFDMGFL-NVAYKKLL
H.inf.    HIYIK-PDRP---XDPDAIKVH---GITDEMLADKPEFKEVAQDFLDYINGAELLIHNAPFDVGFM-DYEFRKLN
E.c.      HVYLK-DRLV----DPEAFGVH---GIAVDFLLDKPTFAEVAVEFMDYIRGAELVIHNAAFDIGFM-DYEFSLLK
H.pyl.    ETLVKVKSVP-----DYIAELT---GITYEDTLNAPSAHEALQELRLFLGNSVFVAHNANFDYNFLGRYFVEKLH

                                                                          3'-Exo IIIC
T.th.     -----YRLENPVVDSLRLARRGLPGLRRYGLDALSEVLELPRRT--CHRALEDVERTLAVVHEVYYMLT-----SG
D.rad.    ----LSWAPERELCTMQLSRRAFPRERTHNLTVLAERLGLEFAPGGRHRSYGDVQVTAQAYLRLLELLG-----ER
Bac.sub.  E---VEKAKNPVIDTLELGRFLYPEFKNHRLNTLCKKFDIELTQ--HHRAIYDTEATAYLLLKMLKDAA-----EK
H.inf.    -LNVKTDDICLVTDTLQMARQMYPGKRN-NLDALCDRLGIDNSKRTLHGALLDAEILADVYLMMTGGQTNLFDEEE
E.c.      RDIAKTNTFCKVTDSLAVARKMFPGKRN-SLDALCARYEIDNSKRTLHGALLDAQILAEVYLAMTGGQTSMAFAME
H.pyl.    -----CPLLNLKLCTLDLSKRAILSMRY-SLSFLKELLGFGIEV--SHRAYADALASYKLFEICLLNLP--SYIKT
```

FIG. 17

FIG.18A

```
ATGGTGGAGCGGGTGGTGCGGACCCTTCTGGACGGGAGGT 40
TCCTCCTGGAGGAGGGGGTGGGGCTTTGGGAGTGGCGCTA
CCCCTTTCCCCTGGAGGGGAGGCGGTGGTGGTCCTGGAC 120
CTGGAGACCACGGGGCTTGCCGGCCTGGACGAGGTGATTG
AGGTGGGCCTCCTCCGCCTGGAGGGGGGGAGGCGCCTCCC 200
CTTCCAGAGCCTCGTCCGGCCCCTCCCGCCCGCCGAAGCC
CGTTCGTGGAACCTCACCGGCATCCCCCGGGAGGCCCTGG 280
AGGAGGCCCCCTCCCTGGAGGAGGTTCTGGAGAAGGCCTA
CCCCCTCCGCGGCGACGCCACCTTGGTGATCCACAACGCC 360
GCCTTTGACCTGGGCTTCCTCCGCCCGGCCTTGGAGGGCC
TGGGCTACCGCCTGGAAAACCCCGTGGTGGACTCCCTGCG 440
CTTGGCCAGACGGGGCTTACCAGGCCTTAGGCGCTACGGC
CTGGACGCCCTCTCCGAGGTCCTGGAGCTTCCCCGAAGGA 520
CCTGCCACCGGGCCCTCGAGGACGTGGAGCGCACCCTCGC
CGTGGTGCACGAGGTATACTATATGCTTACGTCCGGCCGT 600
CCCCGCACGCTTTGGGAACTCGGGAGGTAG
```

```
MVERVVRTLLDGRFLLEEGVGLWEWRYPFPLEGEAVVVLD 40
LETTGLAGLDEVIEVGLLRLEGGRRLPFQSLVRPLPPAEA
RSWNLTGIPREALEEAPSLEEVLEKAYPLRGDATLVIHNA 120
AFDLGFLRPALEGLGYRLENPVVDSLRLARRGLPGLRRYG
LDALSEVLELPRRTCHRALEDVERTLAVVHEVYYMLTSGR 200
PRTLWELGRZ
```

FIG.18B

Alignment of dnaA genes.

```
P.mar.          MLEASWEK VQSSL--KQNLSK-- ----------PSYE TWIRPTEFSG--FKN GELTLIAPNSFSSAW LKNNYSQTIQETAE-   65
Syn.sp.       MVSCENLWQQ ALAIL--ATQLTK-- ----------PAFD TWIKASVLIS--LGD GVATIQVENGFVLNH LQKSYGPLLMEVLT-   67
B.sut.        MENILDLWNQ ALAQI--EKKLSK-- ----------PSFE TWMKSTKAHS--LQG DTLTITAPNEFARDW LESRYLHLIADTIY-   67
M.tub.  MTDDPGSGFTTVWNA VVSELNGDPKVDDGP SSDANLSAPLTPQQR AWLNLVQPLT--IVE GFALLSVPSSFVQNE IERHLRAPITDALS-   87
T.th.         MSHEAVWQH VLEHI--RRSITE-- ----------VEFH TWFERIRPLG--IRD GVLELAVPTSFALDW IRRHYAGLIQEGPR-   66
E.coli        MSLSLWQQ CLARL--QDELPA-- ----------TEFS MWIRPLQAE---LSD NTLALYAPNRFVLDW VRDKYLNNINGLLT-   64
T.mar.          MKER ILQEI--KTRVNR-- ----------KSWE LWFSSFDVKS--IEG NKVVFSVGNLFIKEW LEKKYYSVLSKAVK-   61
H.pyl.        MDTNNNIEKE ILALVKQNPKVSL-- ----------IEYE NYFSQLKYNPNASKS DIAFFYAPNQVLCTT ITAKYGALLKEILSQ   72

P.mar.  EIFG---EPVTVHVK VKANAESSDEHYSSA P---------- ---ITPPLEASPGSV DSSGSSLRLSK---- -KTLPLLNLRYVFNR  130
Syn.sp. DLTG---QEITVKLI TDGLEPHS---LIGQ E---------- ---SSLPMETTP--- ---------------- -KNATALNGKYTFSR  115
B.sut.  ELTG---EELSIKFV IPQNQDVEDFMPKPQ VKKAVKEDTSDFPQN ---------- ---------------- -----MLNPKYTFDT  119
M.tub.  RRLGH-QIQLGVRIA PPATDEADDTTVPPS ENPATTSPDTTTDND EIDDSAAARGDNQHS WPSYFTERPHNTDSA TAGVTSLNRRYTFDT  176
T.th.   LLGAQ-APRFELRVV PGVVVQEDIFQPPPS PPAQAQP-------- ---------- ---------- ---------EDTFKT  108
E.coli  SFCGADAPQLRFEVG TKPVTQTPQAAVTSN VAAPAQVAQTQPQRA APSTRSGWDNVPAPA EP---------- -TYRSNVNVKHTFDN  140
T.mar.  VVLG---NDATFEIT YEAFEPHSSYSEPLV KKRAVLLTP------ ---------- ---------- ------LNPDYTFEN  106
H.pyl.  NKVG-MHLAHSVDVR IEVAPKIQINAQSNI NYKAIKTS------- ---------- ---------- ------VKDSYTFEN  118

P.mar.  FVVGPNSRMAHAAAM AVAESPGREFNPLFI CGGVGLGKTHLMQAI GHYRLEIDPGAKVSY VSTETFTNDLIL--A IRQDRMQAFRDRYR-  217
Syn.sp. FVVGPTNRMAHAASL AVAESPGREFNPLFL CGGVGLGKTHLMQAI AHYRLEMYPNAKVYY VSTERFTNDLIT--A IRQDNMEDFRSYYR-  202
B.sut.  FVIGSGNRFAHAASL AVAEAPAKAYNPLFI YGGVGLGKTHLMHAI GHYVIDHNPSAKVVY LSSEKFTNEFIN--S IRDNKAVDFRNRYR-  206
M.tub.  FVIGASNRFAHAAAL AIAEAPARAYNPLFI WGESGLGKTHLLHAA GNYAQRLFPGMRVKY VSTEEFTNDFIN--S LRDDRKVAFKRSYR-  263
T.th.   SWWGPTTPWPHGGAV AVAESPGRAYNPLFI YGGRGLGKTYLMHAV GPLRAKRFPHMRLEY VSTETFTNELINRPS AR-DRMTEFRERYR-  196
E.coli  FVEGKSNQLARAAAR QVADNPGGAYNPLFL YGGTGLGKTHLLHAV GNGIMARKPNAKVVY MHSERFVQDMVK--A LQNNAIEEFKRYYR-  227
T.mar.  FVVGPGNSFAYHAAL EVAKHPGR-YNPLFI YGGVGLGKTHLLQSI GNYVVQNEPDLRVMY ITSEKFLNDLVD--S MKEGKLNEFREKYRK  193
H.pyl.  FVVGSCNNTVYEIAK KVAQSDTPPYNPVLF YGGTGLGKTHILNAI GNHALEK--HKKVVL VTSEDFLTDFLK--H LDNKTMDSFKAKYR-  203
```

FIG.19A

```
P.mar.   AADLILVDDIQFIEG KEYTQEEFFHTFNAL HDAGSQIVLASDRPP SQIPRLQERLMSRFS MGLIADVQAPDLETR MAILQKKAEHERVGL  307
Syn.sp.  SADFLLIDDIQFIKG KEYTQEEFFHTFNSL HEAGKQVVVASDRAP QRIPGLQDRLISRFS MGLIADIQVPDLETR MAILQKKAEYDRIRL  292
B.sut.   NVDVLLIDDIQFLAG KEQTQEEFFHTFNTL HEESKQIVISSDRPP KEIPTLEDRLRSRFE WGLITDITPPDLETR IAILRKKAKAEGLDI  296
M.tub.   DVDVLLVDDIQFIEG KEGIQEEFFHTFNTL HNANKQIVISSDRPP KQLATLEDRLRTRFE WGLITDVQPPELETR IAILRKKAQMERLAV  353
T.th.    SVDLLLVDDVQFIAG KERTQEEFFHTFNAL YEAHKQIILSSDRPP KDILTLEARLRSRFE WGLITDNPAPDLETR IAILKMNAS-SGPED  285
E.coli   SVDALLIDDIQFFAN KERSQEEFFHTFNAL LEGNQQIILTSDRYP KEINGVEDRLKSRFG WGLTVAIEPPELETR VAILMKKADENDIRL  317
T.mar.   KVDILLIDDVQFLIG KTGVQTELFHTFNEL HDSGKQIVICSDREP QKLSEFQDRLVSRFQ MGLVAKLEPPDEETR KSIARKMLEIEHGEL  283
H.pyl.   HCDFFLLDDAQFLQG KPKLEEEFFHTFNEL HANSKQIVLISDRSP KNIAGLEDRLKSRFE WGITAKVMPPDLETK LSIVKQKCQLNQITL  293

P.mar.   PRDLIQFIAGRFTSN IRELEGALTRAIAFA SITGLPMTVDSIAPM LD----PNGQGVEVT PKQVLDKVAEVFKVT PDEMRSASRRR-PVS  392
Syn.sp.  PKEVIEYIASHYTSN IRELEGALIRAIAYT SLSNVAMTVENIAPV LN----PPVEKVAAA PETIITIVAQHYQLK VEELLSNSRRR-EVS  377
B.sut.   PNEVMLYIANQIDSN IRELEGALIRVVAYS SLINKDINADLAAEA LKDII-PSSKPKVIT IKEIQRVVGQQFNIK LEDFKAKKRTK-SVA  384
M.tub.   PDDVLELIASSIERN IRELEGALIRVTAFA SLNKTPIDKALAEIV LRDLI-ADANTMQIS AATIMAATAEYFDTT VEELRGPGKTR-ALA  441
T.th.    PEDALEYIARQVTSN IREWEGALMRASPFA SLNGVELTRAVAAKA LRHLR-P--RELEAD PLEIIRKAAGPVRPE TPGGAHGERRKKEVV  372
E.coli   PGEVAFFIAKRLRSN VRELEGALNRVIANA NFTGRAITIDFVREA LRDLL-A-LQEKLVT IDNIQKTVAEYYKIK VADLLSKRRSR-SVA  404
T.mar.   PEEVLNFVAENVDDN LRRLRGAIIKLLVYK ETTGKEVDLKEAILL LKDFIKPNRVKAMDP IDELIEIVAKVTGVP REEILSNSRNV-KAL  372
H.pyl.   PEEVMEYIAQHISDN IRQMEGAIIKISVNA NLMNASIDLNLAKTV LEDL--QKDHAEGSS LENILLAVAQSLNLK SSEIKVSSRQK-NVA  380

P.mar.   QARQVGMYLMRQGTN LSLPRIGDTFGGKDH TTVMYAIEQVEKKLS S---------DPQIA SQVQKIRDLLQIDSR RKR----    461
Syn.sp.  LARQVGMYLMRQHTD LSLPRIGEAFGGKDH TTVMYSCDKITQLQQ K---------DWETS QTLTSLSHRINIAGQ APES---    447
B.sut.   FPRQIAMYLSREMTD SSLPKIGEEFGGRDH TTVIHAHEKISKLLA D---------DEQLQ QHVKEIKEQLK---- -------    446
M.tub.   QSRQIAMYLCRELTD LSLPKIGQAFG-RDH TTVMYAQRKILSEMA E---------RREVF DHVKELTTRIRQRSK R------    507
T.th.    LPRQLAMYLVRELTP ASLPEIGQLFGGRDH TTVRYAIQKVQELAG KP--------DREVQ GLLRTLREACTDPVD NLWITCG    446
E.coli   RPRQMAMALAKELTN HSLPEIGDAFGGRDH TTVLHACRKIEQLRE E---------SHDIK EDFSNLIRTLSS--- -------    467
T.mar.   TARRIGMYVAKNYLK SSLRTIAEKFN-RSH PVVVDSVKKVKDSLL KG--------NKQLK ALIDEVIGEISRRAL SG-----    440
H.pyl.   LARKLVVYFARLYTP NPTLSLAQFLDLKDH SSISKMYSGVKKMLE EEKSPFVLSLREEIK NRLNELNDKKTAFNS SE-----    457
```

FIG.19B

GTGTCGCACGAGGCCGTCTGGCAACACGTTCTGGAGCACA
TCCGCCGCAGCATCACCGAGGTGGAGTTCCACACCTGGTT
TGAAAGGATCCGCCCCTTGGGGATCCGGGACGGGGTGCTG 120
GAGCTCGCCGTGCCCACCTCCTTTGCCCTGGACTGGATCC
GGCGCCACTACGCCGGCCTCATCCAGGAGGGCCCTCGGCT
CCTCGGGGCCCAGGCGCCCCGGTTTGAGCTCCGGGTGGTG 240
CCCGGGGTCGTAGTCCAGGAGGACATCTTCCAGCCCCCGC
CGAGCCCCCCGGCCCAAGCTCAACCCGAAGATACCTTTAA
AACTTCGTGGTGGGGCCCAACAACTCCATGGCCCCACGGC 360
GGCGCCGTGGCCGTGGCCGAGTCCCCCGGCCGGGCCTACA
ACCCCCTCTTCATCTACGGGGGCCGTGGCCTGGGAAAGAC
CTACCTGATGCACGCCGTGGGCCCACTCCGTGCGAAGCGC 480
TTCCCCCACATGAGATTAGAGTACGTTTCCACGGAAACTT
TCACCAACGAGCTCATCAACCGGCCATCCGCGAGGGACCG
GATGACGGAGTTCCGGGAGCGGTACCGCTCCGTGGACCTC 600
CTGCTGGTGGACGACGTCCAGTTCATCGCCGGAAAGGAGC
GCACCCAGGAGGAGTTTTTCCACACCTTCAACGCCCTTTA
CGAGGCCCACAAGCAGATCATCCTCTCCTCCGACCGGCCG 720
CCCAAGGACATCCTCACCCTGGAGGCGCGCCTGCGGAGCC
GCTTTGAGTGGGGCCTGATCACCGACAATCCAGCCCCCGA
CCTGGAAACCCGGATCGCCATCCTGAAGATGAACGCCAGC 840
AGCGGGCCTGAGGATCCCGAGGACGCCCTGGAGTACATCG
CCCGGCAGGTCACCTCCAACATCCGGGAGTGGGAAGGGGC
CCTCATGCGGGCATCGCCTTTCGCCTCCCTCAACGGCGTT 960
GAGCTGACCCGCGCCGTGGCGGCCAAGGCTCTCCGACATC
TTCGCCCCAGGGAGCTGGAGGCGGACCCCTTGGAGATCAT
CCGCAAAGCGGCGGGACCAGTTCGGCCTGAAACCCCGGGA 1080
GGAGCTCACGGGGAGCGCCGCAAGAAGGAGGTGGTCCTCC
CCCGGCAGCTCGCCATGTACCTGGTGCGGGAGCTCACCCC
GGCCTCCCTGCCCGAGATCGACCAGCTCAACGACGACCGG 1200
GACCACACCACGGTCCTCTACGCCATCCAGAAGGTCCAGG
AGCTCGCGGAAAGCGACCGGGAGGTGCAGGGCCTCCTCCG
CACCCTCCGGGAGGCGTGCACATGA

FIG.20A

VSHEAVWQHVLEHIRRSITEVEFHTWFERIRPLGIRDGVL
ELAVPTSFALDWIRRHYAGLIQEGPRLLGAQAPRFELRVV
PGVVVQEDIFQPPPSPPAQAQPEDTFKTSWWGPTTPWPHG 120
GAVAVAESPGRAYNPLFIYGGRGLGKTYLMHAVGPLRAKR
FPHMRLEYVSTETFTNELINRPSARDRMTEFRERYRSVDL
LLVDDVQFIAGKERTQEEFFHTFNALYEAHKQIILSSDRP 240
PKDILTLEARLRSRFEWGLITDNPAPDLETRIAILKMNAS
SGPEDPEDALEYIARQVTSNIREWEGALMRASPFASLNGV
ELTRAVAAKALRHLRPRELEADPLEIIRKAAGPVRPETPG 360
GAHGERRKKEVVLPRQLAMYLVRELTPASLPEIDQLNDDR
DHTTVLYAIQKVQELAESDREVQGLLRTLREACT

FIG.20B

```
ATGAACATAACGGTTCCCAAAAAACTCCTCTCGGACCAGC   40
TTTCCCTCCTGGAGCGCATCGTCCCCTCTAGAAGCGCCAA
CCCCCTCTACACCTACCTGGGGCTTTACGCCGAGGAAGGG  120
GCCTTGATCCTCTTCGGGACCAACGGGGAGGTGGACCTCG
AGGTCCGCCTCCCCGCCGAGGCCCAAAGCCTTCCCCGGGT  200
GCTCGTCCCCGCCCAGCCCTTCTTCCAGCTGGTGCGGAGC
CTTCCTGGGGACCTCGTGGCCCTCGGCCTCGCCTCGGAGC  280
CGGGCCAGGGGGGGCAGCTGGAGCTCTCCTCCGGGCGTTT
CCGCACCCGGCTCAGCCTGGCCCCTGCCGAGGGCTACCCC  360
GAGCTTCTGGTGCCCGAGGGGGAGGACAAGGGGGCCTTCC
CCCTCCGGACGCGGATGCCCTCCGGGGAGCTCGTCAAGGC  440
CTTGACCCACGTGCGCTACGCCGCGAGCAACGAGGAGTAC
CGGGCCATCTTCCGCGGGGTGCAGCTGGAGTTCTCCCCCC  520
AGGGCTTCCGGGCGGTGGCCTCCGACGGGTACCGCCTCGC
CCTCTACGACCTGCCCCTGCCCCAAGGGTTCCAGGCCAAG  600
GCCGTGGTCCCCGCCCGGAGCGTGGACGAGATGGTGCGGG
TCCTGAAGGGGGCGGACGGGGCCGAGGCCGTCCTCGCCCT  680
GGGCGAGGGGGTGTTGGCCCTGGCCCTCGAGGGCGGAAGC
GGGGTCCGGATGGCCCTCCGCCTCATGGAAGGGGAGTTCC  760
CCGACTACCAGAGGGTCATCCCCCAGGAGTTCGCCCTCAA
GGTCCAGGTGGAGGGGGAGGCCCTCAGGGAGGCGGTGCGC  840
CGGGTGAGCGTCCTCTCCGACCGGCAGAACCACCGGGTGG
ACCTCCTTTTGGAGGAAGGCCGGATCCTCCTCTCCGCCGA  920
GGGGGACTACGGCAAGGGGCAGGAGGAGGTGCCCGCCCAG
GTGGAGGGGCCGGACATGGCCGTGGCCTACAACGCCCGCT 1000
ACCTCCTCGAGGCCCTCGCCCCCGTGGGGGACCGGGCCCA
CCTGGGCATCTCCGGGCCCACGAGCCCGAGCCTCATCTGG 1080
GGGGACGGGGAGGGGTACCGGGCGGTGGTGGTGCCCCTCA
GGGTCTAG                                  1128
```

FIG.21A

```
MNITVPKKLLSDQLSLLERIVPSRSANPLYTYLGLYAEEG 40
ALILFGTNGEVDLEVRLPAEAQSLPRVLVPAQPFFQLVRS
LPGDLVALGLASEPGQGGQLELSSGRFRTRLSLAPAEGYP 120
ELLVPEGEDKGAFPLRTRMPSGELVKALTHVRYAASNEEY
RAIFRGVQLEFSPQGFRAVASDGYRLALYDLPLPQGFQAK 200
AVVPARSVDEMVRVLKGADGAEAVLALGEGVLALALEGGS
GVRMALRLMEGEFPDYQRVIPQEFALKVQVEGEALREAVR 280
RVSVLSDRQNHRVDLLLEEGRILLSAEGDYGKGQEEVPAQ
VEGPDMAVAYNARYLLEALAPVGDRAHLGISGPTSPSLIW 360
GDGEGYRAVVVPLRVZ
```

FIG.21B

```
T.th.beta       MNITVPKKLLSDQLSLLERIVPSRSANPLYTYLGLYAEEGALILFGTNGEVDLEVRLPAE
E.coli.bet      MKFTVEREHLLKPLQQVSGPLGGRPTLPILGNLLLQVADGTLSLTGTDLEMEMVARVALV
P.mirab.be      MKFIIEREQLLKPLQQVSGPLGGRPTLPILGNLLLKVTENTLSLTGTDLEMEMMARVSLS
H.infl.bet      MQFSISRENLLKPLQQVCGVLSNRPNIPVLNNVLLQIEDYRLTITGTDLEVELSSQTQLS
P.put.beta      MHFTIQREALLKPLQLVAGVVERRQTLPVLSNVLLVVQGQQLSLTGTDLEVELVGRVQLE
B.cap.beta      MKFTIQNDILTKNLKKITRVLVKNISFPILENILIQVEDGTLSLTTTNLEIELISKIEII
                *   .   *  *  .  .      *. . .    *  . *  *... .

T.th.beta       AQSLP-RVLVPAQPFFQLVRSLPGDLVALGLASEPGQGGQLELSSGRFRTRLSLAPAEGY
E.coli.bet      QPHEPGATTVPARKFFDICRGLP-EGAEIAVQLE---GERMLVRSGRSRFSLSTLPAADF
P.mirab.be      QSHEIGATTVPARKFFDIWRGLP-EGAEISVELD---GDRLLVRSGRSRFSLSTLPASDF
H.infl.bet      SSSENGTFTIPAKKFLDICRTLS-DDSEITVTFE---QDRALVQSGRSRFTLATQPAEEY
P.put.beta      EPAEPGEITVPARKLMDICKSLP-NDALIDIKVD---EQKLLVKAGRSRFTLSTLPANDF
B.cap.beta      TKYIPGKTTISGRKILNICRTLS-EKSKIKMQLK---NKKMYISSENSNYILSTLSADTF
                        . .   . .*     . .       . . .      *.  *  .

T.th.beta       PELLVPEGEDKGAFPLRTRMPSGELVKALTHVRYAASNEEYRAIFRGVQLEFSPQGFRAV
E.coli.bet      PNLDD--WQSEVEFTLPQAT----MKRLIEATQFSMAHQDVRYYLNGMLFETEGEELRTV
P.mirab.be      PNLDD--WQSEVEFTLPQAT----LKRLIESTQFSMAHQDVRYYLNGMLFETENTELRTV
H.infl.bet      PNLTD--WQSEVDFELPQNT----LRRLIEATQFSMANQDARYFLNGMKFETEGNLLRTV
P.put.beta      PTVEE--GPGSLTCNLEQSK----LRRLIERTSFAMAQQDVRYYLNGMLLEVSRNTLRAV
B.cap.beta      PNHQN--FDYISKFDISSNI----LKEMIEKTEFSMGKQDVRYYLNGMLLEKKDKFLRSV
                *                .   .  .  ..  ..*   *.  *     *.*

T.th.beta       ASDGYRLALYDLPLPQGFQA--KAVVPARSVDEMVRVLKGADGAEAVLALGEGVLALALE
E.coli.bet      ATDGHRLAVCSMPIGQSLPS-HSVIVPRKGVIELMRMLDG-GDNPLRVQIGSNNIRAHVG
P.mirab.be      ATDGHRLAVCAMDIGQSLPG-HSVIVPRKGVIELMRLLDGSGESLLQLQIGSNNLRAHVG
H.infl.bet      ATDGHRLAVCTISLEQELQN-HSVILPRKGVLELVRLLET-NDEPARLQIGTNNLRVHLK
P.put.beta      STDGHRLALCSMSAPIEQEDRHQVIVPRKGILELARLLTD-PEGMVSIVLGQHHIRATTG
B.cap.beta      ATDGYRLAISYTQLKKDINF-FSIIIPNKAVMELLKLLNT-QPQLLNILIGSNSIRIYTK
                ..** ***. .          ..* . . *. ..*        . .*   .
```

FIG.22A

```
T.th.beta       GGSGVRMALRLMEGEFPDYQRVIPQEFALKVQVEGEALREAVRRVSVLSDRQNHRVDLLL
E.coli.bet      ---DFIFTSKLVDGRFPDYRRVLPKNPDKHLEAGCDLLKQAFARAAILSNEKFRGVRLYV
P.mirab.be      ---DFIFTSKLVDGRFPDYRRVLPKNPTKTVIAGCDILKQAFSRAAILSNEKFRGVRINL
H.infl.bet      ---NTVFTSKLIDGRFPDYRRVLPRNATKIVEGNWEMLKQAFARASILSNERARSVRLSL
P.put.beta      ---EFTFTSKLVDGKFPDYERVLPKGGDKLVVGDRQALREAFSRTAILSNEKYRGIRLQL
B.cap.beta      ---NLIFTTQLIEGEYPDYKSVLFKEKKNPIITNSILLKKSLLRVAILAHEKFCGIEIKI
                   . .*..* .***  *. .    .      *. .  * ..*.  .   . . .

T.th.beta       EEGRILLSAEGDYGK-GQEEVPAQVEGPDMAVAYNARYLLEALAPVG-DRAHLGISGPTS
E.coli.bet      SENQLKITANNPEQEEAEEILDVTYSGAEMEIGFNVSYVLDVLNALKCENVRMMLTDSVS
P.mirab.be      TNGQLKITANNPEQEEAEEIVDVQYQGEEMEIGFNVSYLLDVLNTLKCEEVKLLLTDAVS
H.infl.bet      KENQLKITASNTEHEEAEEIVDVNYNGEELEVGFNVTYILDVLNALKCNQVRMCLTDAFS
P.put.beta      AAGQLKIQANNPEQEEAEEEISVDYEGSSLEIGFNVSYLLDVLGVMITEQVRLILSDSNS
B.cap.beta      ENGKFKVLSDNQEEETAEDLFEIDYFGEKIEISINVYYLLDVINNIKSENIALFLNKSKS
                  ..  . . .     ..        *  . .  *  *.*. . .  .      . .   *

T.th.beta       PSLIWGDG-EGYRAVVVPLRVZ      (ID#108)
E.coli.bet      SVQIEDAASQSAAYVVMPMRLZ      (ID#109)
P.mirab.be      SVQVENVASAAAAYVVMPMRL-      (ID#110)
H.infl.bet      SCLIENCEDSSCEYVIMPMRL-      (ID#111)
P.put.beta      SALLQEAGNDDSSYVVMPMRL-      (ID#112)
B.cap.beta      SIQIEAENNSSNAYVVMLLKR-      (ID#113)
                   .        *.. ..
```

FIG.22B

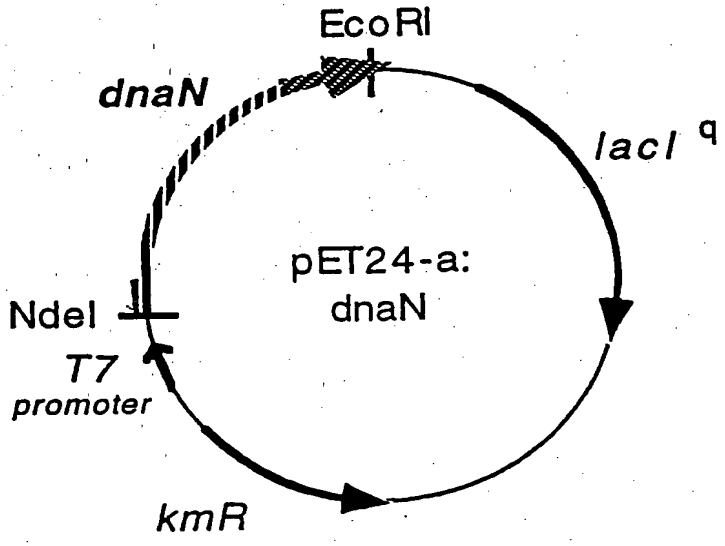


FIG.23

FIG.24A

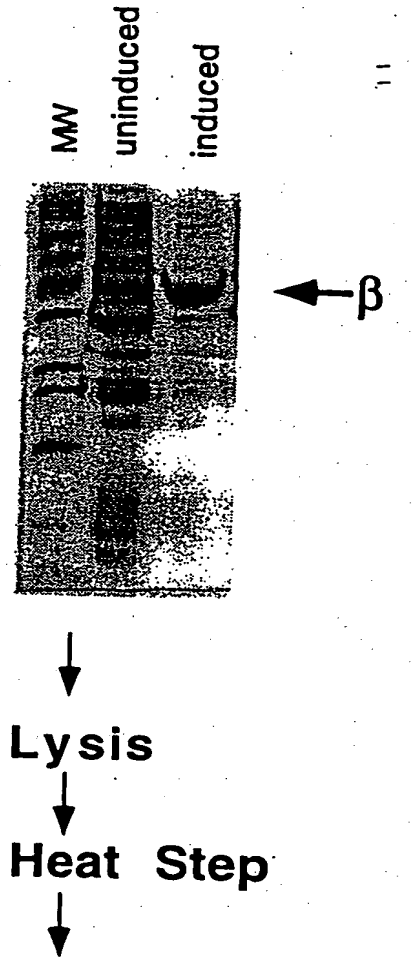


FIG.24B

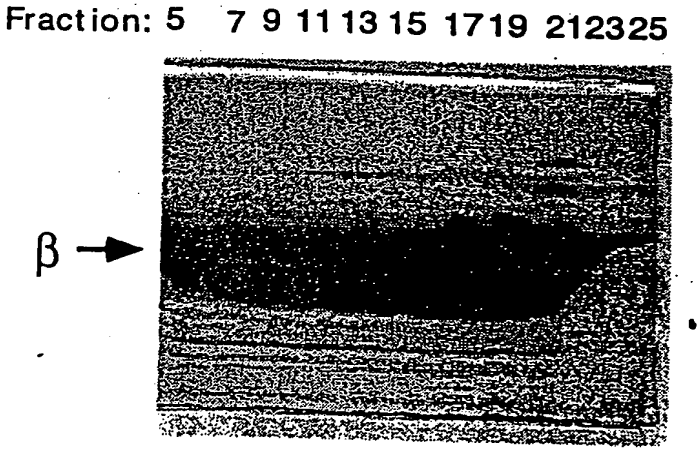
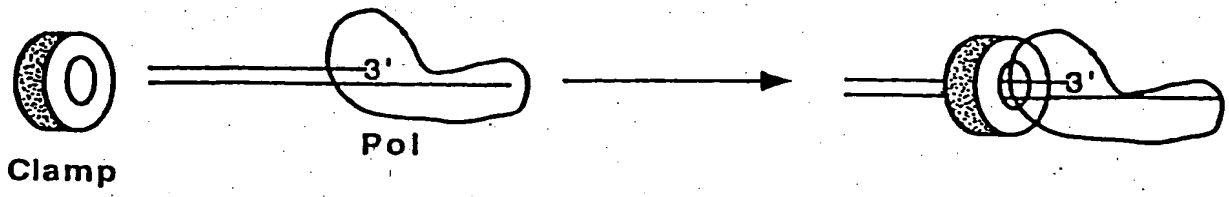


FIG.25A



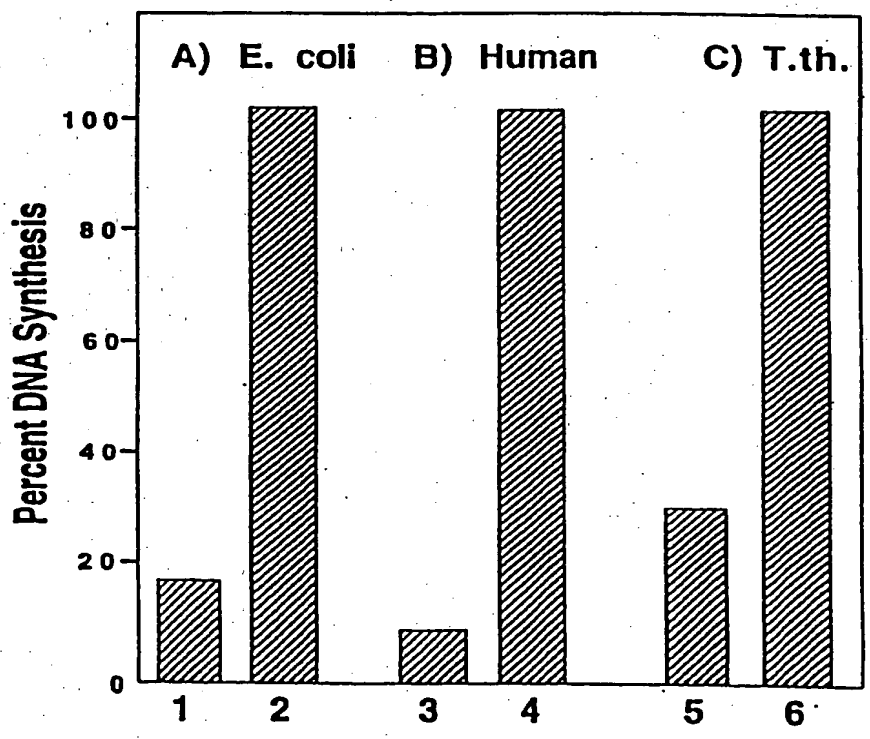
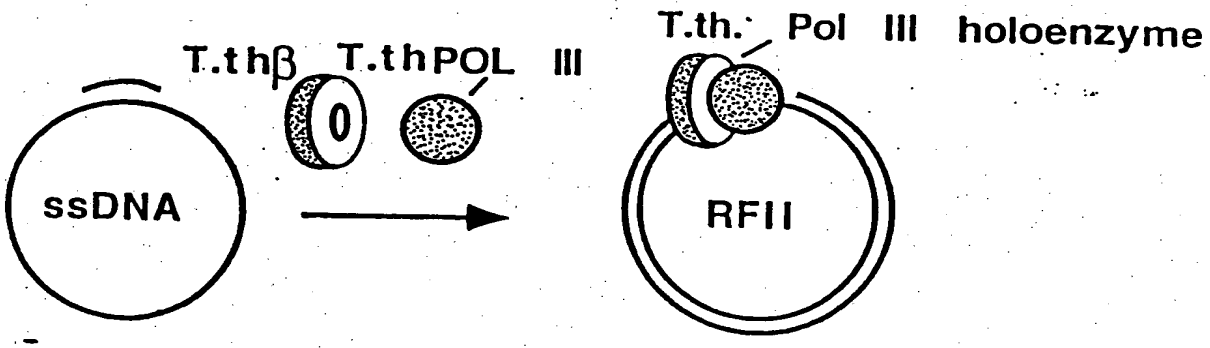


FIG.25B

FIG. 26A



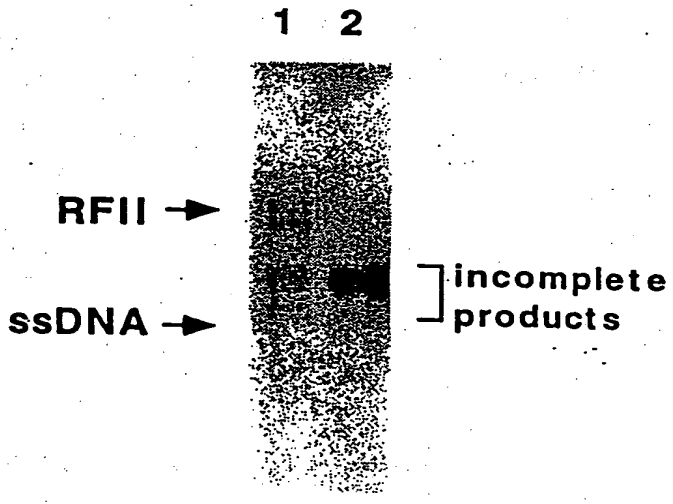


FIG. 26B

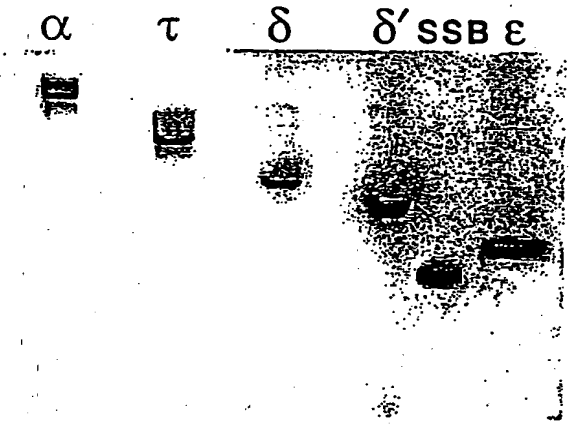

α
τ
δ
δ′
SSB
ε

FIG. 27

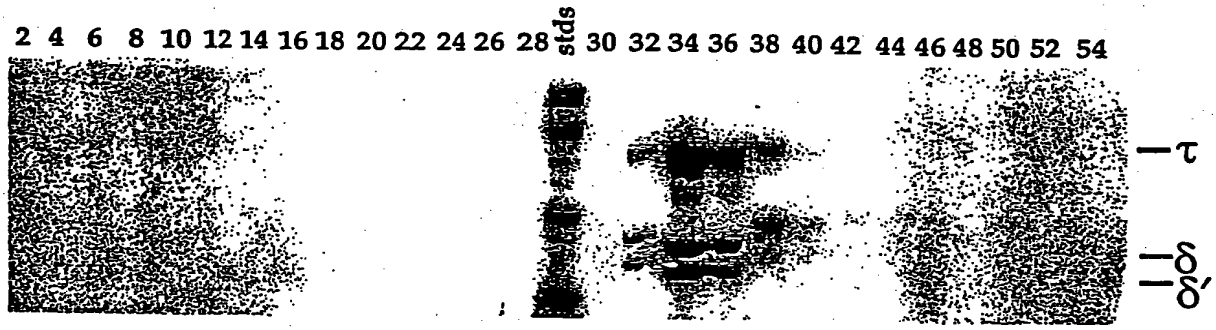

2 4 6 8 10 12 14 16 18 20 22 24 26 28
stds
30 32 34 36 38 40 42 44 46 48 50 52 54
τ
δ
δ′

FIG. 28

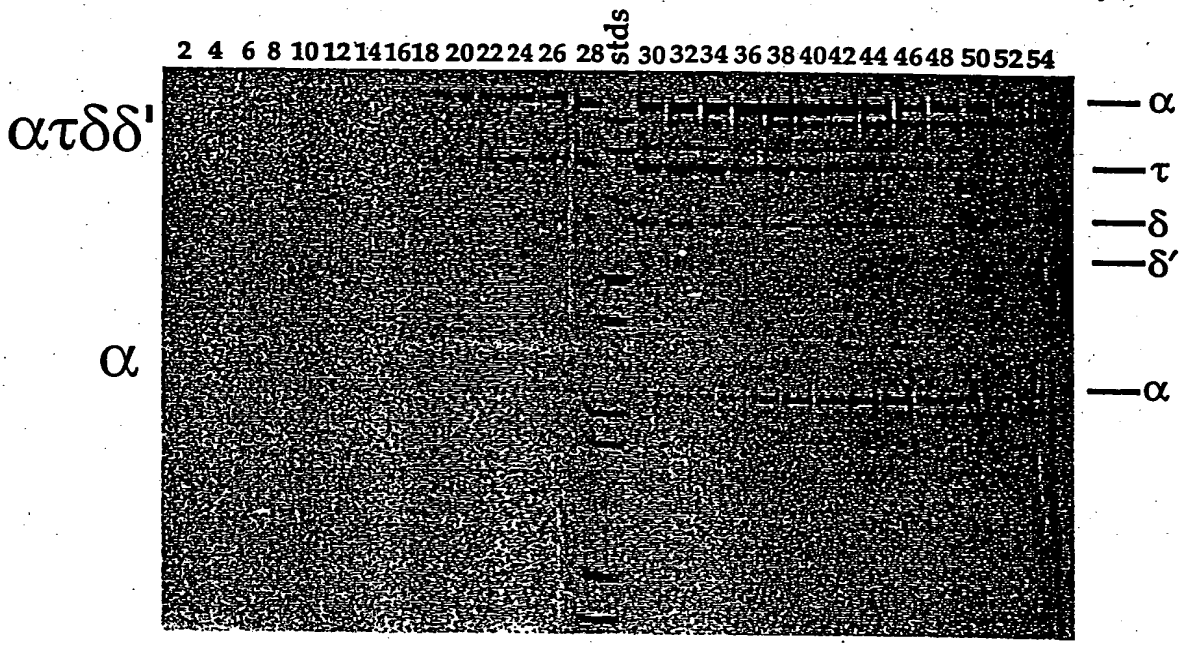

2 4 6 8 10 12 14 16 18 20 22 24 26 28 stds 30 32 34 36 38 40 42 44 46 48 50 52 54
αταδδ'
α
α
τ
δ
δ'
α

FIG. 29

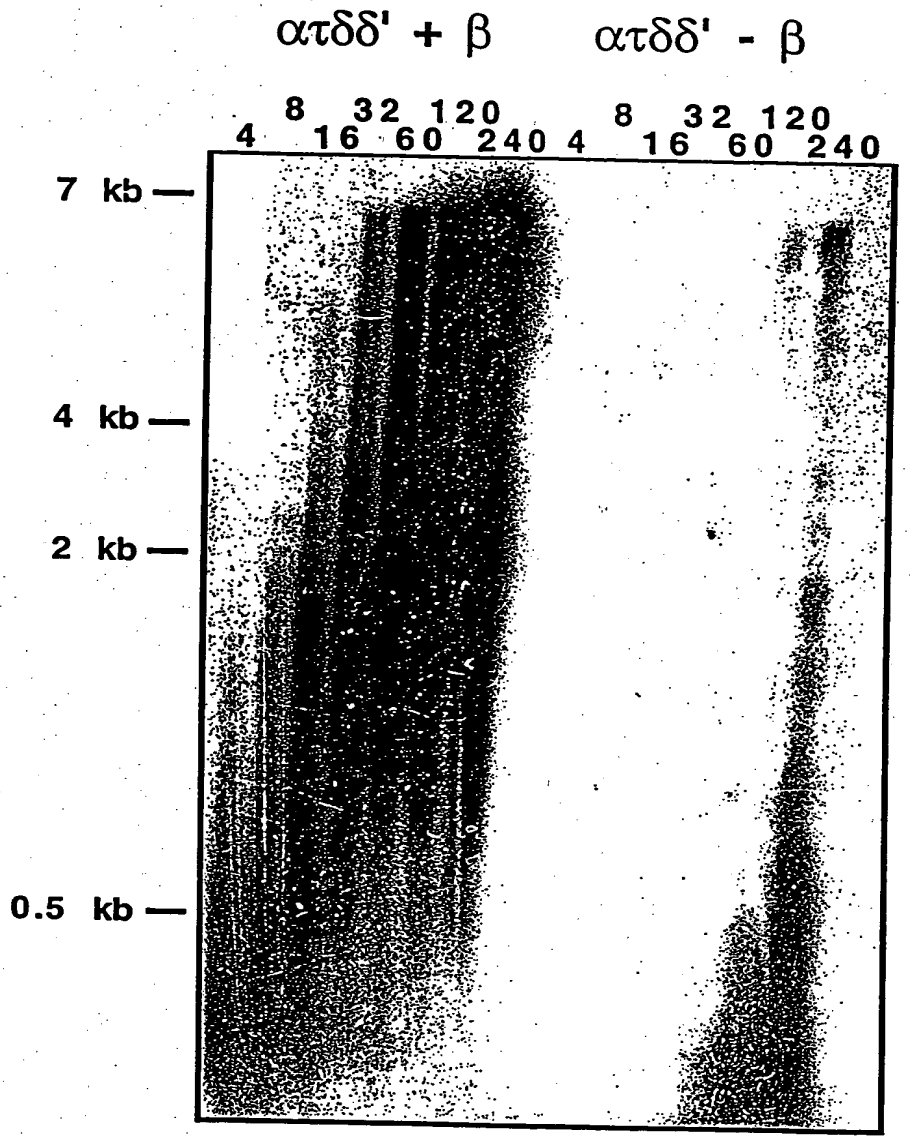

ατδδ' + β
ατδδ' - β
4 8 16 32 60 120 240
4 8 16 32 60 120 240
7 kb
4 kb
2 kb
0.5 kb

FIG. 30

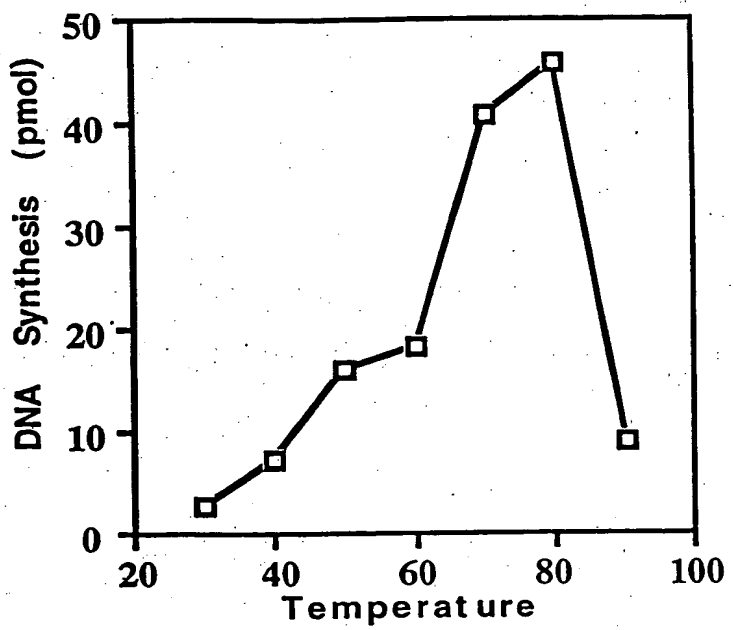

DNA Synthesis (pmol)
50
40
30
20
10
0
20
40
60
80
100
Temperature

FIG. 31

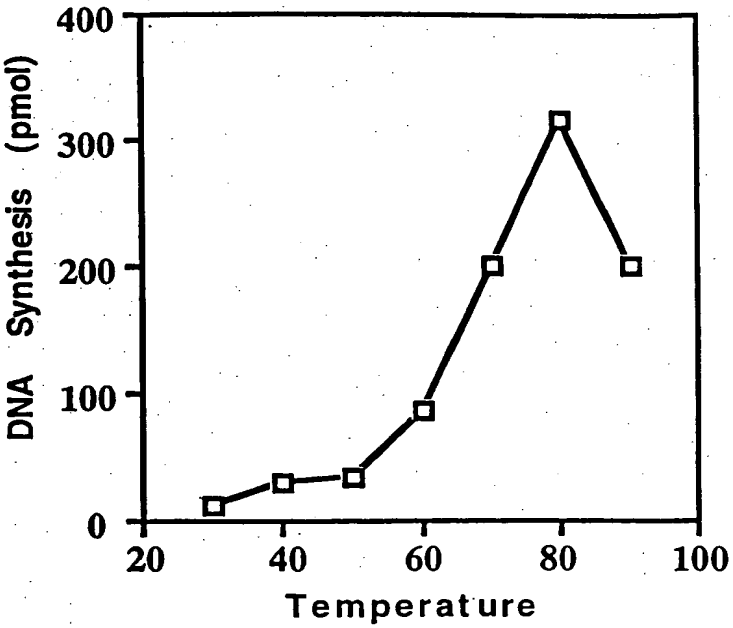

DNA Synthesis (pmol)
400
300
200
100
0
20
40
60
80
100
Temperature

FIG. 32

α

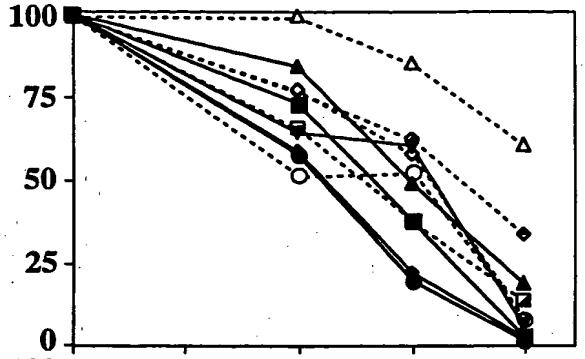


FIG. 33A

β

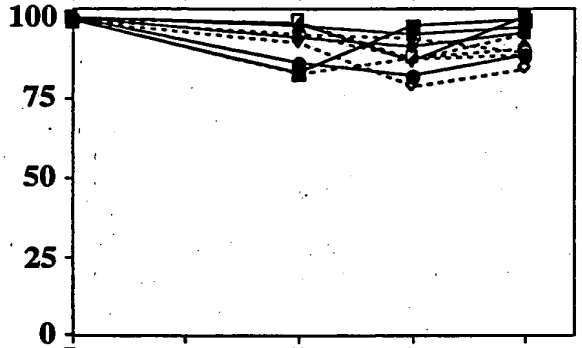


FIG. 33B

τδδ'

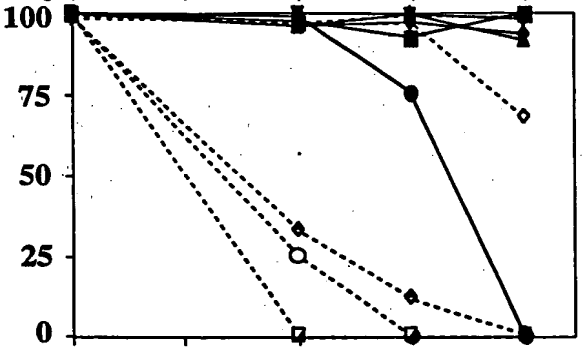


FIG. 33C

SSB

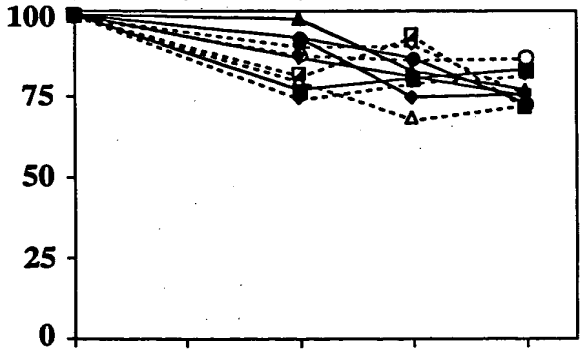


FIG. 33D

Pol III*

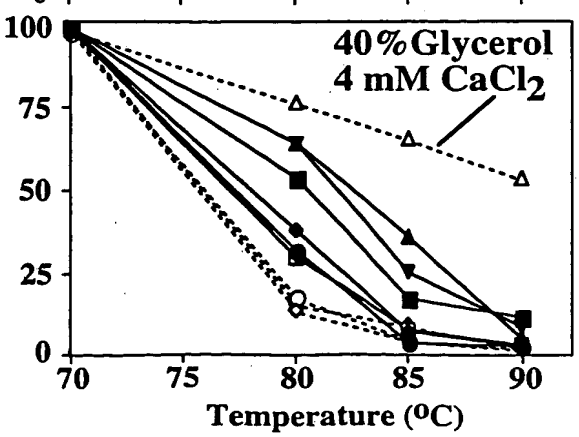


FIG. 33E

Activity (%)

ATGAGTAAGGATTTCGTCCACCTTCACCTGCACACCCAGTTCTCACTCCT
GGACGGGGCTATAAAGATAGACGAGCTCGTGAAAAAGGCAAAGGAGTATG 100
GATACAAAGCTGTCGGAATGTCAGACCACGGAAACCTCTTCGGTTCGTAT
AAATTCTACAAAGCCCTGAAGGCGGAAGGAATTAAGCCCATAATCGGCAT 200
GGAAGCCTACTTTACCACGGGTTCGAGGTTTGACAGAAAGACTAAAACGA
GCGAGGACAACATAACCGACAAGTACAACCACCACCTCATACTTATAGCA 300
AAGGACGAAAAGGTCTAAAGAACTTAATGAAGCTCTCAACCCTCGCCTAC
AAAGAAGGTTTTTACTACAAACCCAGAATTGATTACGAACTCCTTGAAAA 400
GTACGGGGAGGGCCTAATAGCCCTTACCGCATGCCTGAAAGGTGTTCCCA
CCTACTACGCTTCTATAAACGAAGTGAAAAAGGCGGAGGAATGGGTAAAG 500
AAGTTCAAGGATATATTCGGAGATGACCTTTATTTAGAACTTCAAGCGAA
CAACATTCCAGAACAGGAAGTGGCAAACAGGAACTTAATAGAGATAGCCA 600
AAAAGTACGATGTGAAACTCATAGCGACGCAGGACGCCCACTACCTCAAT
CCCGAAGACAGGTACGCCCACACGGTTCTTATGGCACTTCAAATGAAAAA 700
GACCATTCACGAACTGAGTTCGGGAAACTTCAAGTGTTCAAACGAAGACC
TTCACTTTGCTCCACCCGAGTACATGTGGAAAAAGTTTGAAGGTAAGTTC 800
GAAGGCTGGGAAAAGGCACTCCTGAACACTCTCGAGGTAATGGAAAAGAC
AGCGGACAGCTTTGAGATATTTGAAAACTCCACCTACCTCCTTCCCAAGT 900
ACGACGTTCCGCCCGACAAAACCCTTGAGGAATACCTCAGAGAACTCGCG
TACAAAGGTTTAAGACAGAGGATAGAAAGGGGACAAGCTAAGGATACTAA 1000
AGAGTACTGGGAGAGGCTCGAGTACGAACTGGAAGTTATAAACAAAATGG
GCTTTGCGGGATACTTCTTGATAGTTCAGGACTTCATAAACTGGGCTAAG 1100
AAAAACGACATACCTGTTGGACCCGGAAGGGGAAGTGCTGGAGGTTCCCT
CGTCGCATACGCCATCGGAATAACGGACGTTGACCCTATAAAGCACGGAT 1200
TCCTTTTTGAGAGGTTCTTAAACCCCGAAAGGGTTTCCATGCCGGATATA
GACGTGGATTTCTGTCAGGACAACAGGGAAAAGGTCATAGAGTACGTAAG 1300
GAACAAGTACGGACACGACAACGTAGCTCAGATAATCACCTACAACGTAA
TGAAGGCGAAGCAAACACTGAGAGACGTCGCAAGGGCCATGGGACTCCCC 1400
TACTCCACCGCGGACAAACTCGCAAAACTCATTCCTCAGGGGGACGTTCA
GGGAACGTGGCTCAGTCTGGAAGAGATGTACAAAACGCCTGTGGAGGAAC 1500
TCCTTCAGAAGTACGGAGAACACAGAACGGACATAGAGGACAACGTAAAG
AAGTTCAGACAGATATGCGAAGAAAGTCCGGAGATAAAACAGCTCGTTGA 1600
GACGGCCCTGAAGCTTGAAGGTCTCACGAGACACACCTCCCTCCACGCCG
CGGGAGTGGTTATAGCACCAAAGCCCTTGAGCGAGCTCGTTCCCCTCTAC 1700
TACGATAAAGAGGGCGAAGTCGCAACCCAGTACGACATGGTTCAGCTCGA
AGAACTCGGTCTCCTGAAGATGGACTTCCTCGGACTCAAAACCCTCACAG 1800
AACTGAAACTCATGAAAGAACTCATAAAGGAAAGACACGGAGTGGATATA
AACTTCCTTGAACTTCCCCTTGACGACCCGAAAGTTTACAAACTCCTTCA 1900
GGAAGGAAAAACCACGGGAGTGTTCCAGCTCGAAAGCAGGGGAATGAAAG
AACTCCTGAAGAAACTAAAGCCCGACAGCTTTGACGACATCGTTGCGGTC 2000
CTCGCACTCTACAGACCCGGACCTCTAAAGAGCGGACTCGTTGACACATA
CATTAAGAGAAAGCACGGAAAAGAACCCGTTGAGTACCCCTTCCCGGAGC 2100
TTGAACCCGTCCTTAAGGAAACCTACGGAGTAATCGTTTATCAGGAACAG
GTGATGAAGATGTCTCAGATACTTTCCGGCTTTACTCCCGGAGAGGCGGA 2200
TACCCTCAGAAAGGCGATAGGTAAGAAGAAAGCGGATTTAATGGCTCAGA
TGAAAGACAAGTTCATACAGGGAGCGGTGGAAAGGGGATACCCTGAAGAA 2300
AAGATAAGGAAGCTCTGGGAAGACATAGAGAAGTTCGCTTCCTACTCCTT
CAACAAGTCTCACTCGGTAGCTTACGGGTACATCTCCTACTGGACCGCCT 2400

FIG. 34A

ACGTTAAAGCCCACTATCCCGCGGAGTTCTTCGCGGTAAAACTCACAACT
GAAAAGAACGACAACAAGTTCCTCAACCTCATAAAAGACGCTAAACTCTT 2500
CGGATTTGAGATACTTCCCCCCGACATAAACAAGAGTGATGTAGGATTTA
CGATAGAAGGTGAAAACAGGATAAGGTTCGGGCTTGCGAGGATAAAGGGA 2600
GTGGGAGAGGAAACTGCTAAGATAATCGTTGAAGCTAGAAAGAAGTATAA
GCAGTTCAAAGGGCTTGCGGACTTCATAAACAAAACCAAGAACAGGAAGA 2700
TAAACAAGAAAGTCGTGGAAGCACTCGTAAAGGCAGGGGCTTTTGACTTT
ACTAAGAAAAAGAGGAAAGAACTACTCGCTAAAGTGGCAAACTCTGAAAA 2800
AGCATTAATGGCTACACAAAACTCCCTTTTCGGTGCACCGAAAGAAGAAG
TGGAAGAACTCGACCCCTTAAAGCTTGAAAAGGAAGTTCTCGGTTTTTAC 2900
ATTTCAGGGCACCCCCTTGACAACTACGAAAAGCTCCTCAAGAACCGCTA
CACACCCATTGAAGATTTAGAAGAGTGGGACAAGGAAAGCGAAGCGGTGC 3000
TTACAGGAGTTATCACGGAACTCAAAGTAAAAAAGACGAAAAACGGAGAT
TACATGGCGGTCTTCAACCTCGTTGACAAGACGGGACTAATAGAGTGTGT 3100
CGTCTTCCCGGGAGTTTACGAAGAGGCAAAGGAACTGATAGAAGAGGACA
GAGTAGTGGTAGTCAAAGGTTTTCTGGACGAGGACCTTGAAACGGAAAAT 3200
GTCAAGTTCGTGGTGAAAGAGGTTTTCTCCCCTGAGGAGTTCGCAAAGGA
GATGAGGAATACCCTTTATATATTCTTAAAAAGAGAGCAAGCCCTAAACG 3300
GCGTTGCCGAAAAACTAAAGGGAATTATTGAAAACAACAGGACGGAGGAC
GGATACAACTTGGTTCTCACGGTTGATCTGGGAGACTACTTCGTTGATTT 3400
AGCACTCCCACAAGATATGAAACTAAAGGCTGACAGAAAGGTTGTAGAGG
AGATAGAAAAACTGGGAGTGAAGGTCATAATTTAGTAAATAACCCTTACT 3500
TCCGAGTAGTCCCC

FIG. 34B

```
MSKDFVHLHLHTQFSLLDGAIKIDELVKKAKEYGYKAVGMSDHGNLFGSY
KFYKALKAEGIKPIIGMEAYFTTGSRFDRKTKTSEDNITDKYNHHLILIA   100
KDDKGLKNLMKLSTLAYKEGFYYKPRIDYELLEKYGEGLIALTACLKGVP
TYYASINEVKKAEEWVKKFKDIFGDDLYLELQANNIPEQEVANRNLIEIA   200
KKYDVKLIATQDAHYLNPEDRYAHTVLMALQMKKTIHELSSGNFKCSNED
LHFAPPEYMWKKFEGKFEGWEKALLNTLEVMEKTADSFEIFENSTYLLPK   300
YDVPPDKTLEEYLRELAYKGLRQRIERGQAKDTKEYWERLEYELEVINKM
GFAGYFLIVQDFINWAKKNDIPVGPGRGSAGGSLVAYAIGITDVDPIKHG   400
FLFERFLNPERVSMPDIDVDFCQDNREKVIEYVRNKYGHDNVAQIITYNV
MKAKQTLRDVARAMGLPYSTADKLAKLIPQGDVQGTWLSLEEMYKTPVEE   500
LLQKYGEHRTDIEDNVKKFRQICEESPEIKQLVETALKLEGLTRHTSLHA
AGVVIAPKPLSELVPLYYDKEGEVATQYDMVQLEELGLLKMDFLGLKTLT   600
ELKLMKELIKERHGVDINFLELPLDDPKVYKLLQEGKTTGVFQLESRGMK
ELLKKLKPDSFDDIVAVLALYRPGPLKSGLVDTYIKRKHGKEPVEYPFPE   700
LEPVLKETYGVIVYQEQVMKMSQILSGFTPGEADTLRKAIGKKKADLMAQ
MKDKFIQGAVERGYPEEKIRKLWEDIEKFASYSFNKSHSVAYGYISYWTA   800
YVKAHYPAEFFAVKLTTEKNDNKFLNLIKDAKLFGFEILPPDINKSDVGF
TIEGENRIRFGLARIKGVGEETAKIIVEARKKYKQFKGLADFINKTKNRK   900
INKKVVEALVKAGAFDFTKKKRKELLAKVANSEKALMATQNSLFGAPKEE
VEELDPLKLEKEVLGFYISGHPLDNYEKLLKNRYTPIEDLEEWDKESEAV  1000
LTGVITELKVKKTKNGDYMAVFNLVDKTGLIECVVFPGVYEEAKELIEED
RVVVVKGFLDEDLETENVKFVVKEVFSPEEFAKEMRNTLYIFLKREQALN  1100
GVAEKLKGIIENNRTEDGYNLVLTVDLGDYFVDLALPQDMKLKADRKVVE
EIEKLGVKVII                                          1161
```

FIG. 35

ATGAACTACGTTCCCTTCGCGAGAAAGTACAGACCGAAATTCTTCAGGGA
AGTAATAGGACAGGAAGCTCCCGTAAGGATACTCAAAAACGCTATAAAAA 100
ACGACAGAGTGGCTCACGCCTACCTCTTTGCCGGACCGAGGGGGGTTGGG
AAGACGACTATTGCAAGAATTCTCGCAAAAGCTTTGAACTGTAAAAATCC 200
CTCCAAAGGTGAGCCCTGCGGTGAGTGCGAAAACTGCAGGGAGATAGACA
GGGGTGTGTTCCCTGACTTAATTGAAATGGATGCCGCCTCAAACAGGGGT 300
ATAGACGACGTAAGGGCATTAAAAGAAGCGGTCAATTACAAACCTATAAA
AGGAAAGTACAAGGTTTACATAATAGACGAAGCTCACATGCTCACGAAAG 400
AAGCTTTCAACGCTCTCTTAAAAACCCTCGAAGAGCCCCCTCCCAGAACT
GTTTTCGTCCTTTGTACCACGGAGTACGACAAAATTCTTCCCACGATACT 500
CTCAAGGTGTCAGAGGATAATCTTCTCAAAGGTAAGAAAGGAAAAAGTAA
TAGAGTATCTAAAAAAGATATGTGAAAAGGAAGGGATTGAGTGCGAAGAG 600
GGAGCCCTTGAGGTTCTGGCTCATGCCTCTGAAGGGTGCATGAGGGATGC
AGCCTCTCTCCTGGACCAGGCGAGCGTTTACGGGGAAGGCAGGGTAACAA 700
AAGAAGTAGTGGAGAACTTCCTCGGAATTCTCAGTCAGGAAAGCGTTAGG
AGTTTTCTGAAATTGCTTCTGAACTCAGAAGTGGACGAAGCTATAAAGTT 800
CCTCAGAGAACTCTCAGAAAAGGGCTACAACCTGACCAAGTTTTGGGAGA
TGTTAGAAGAGGAAGTGAGAAACGCAATTTTAGTAAAGAGCCTGAAAAAT 900
CCCGAAAGCGTGGTTCAGAACTGGCAGGATTACGAAGACTTCAAAGACTA
CCCTCTGGAAGCCCTCCTCTACGTTGAGAACCTGATAAACAGGGGTAAAG 1000
TTGAAGCGAGAACGAGAGAACCCTTAAGAGCCTTTGAACTCGCGGTAATA
AAGAGCCTTATAGTCAAAGACATAATTCCCGTATCCCAGCTCGGAAGTGT 1100
GGTAAAGGAAACCAAAAAGGAAGAAAAGAAAGTTGAAGTAAAAGAAGAGC
CAAAAGTAAAAGAAGAAAAACCAAAGGAGCAGGAAGAGGACAGGTTCCAG 1200
AAAGTTTTAAACGCTGTGGACGGCAAAATCCTTAAAAGAATACTTGAAGG
GGCAAAAAGGGAAGAAAGAGACGGAAAAATCGTCCTAAAGATAGAAGCCT 1300
CTTATCTGAGAACCATGAAAAAGGAATTTGACTCACTAAAGGAGACTTTT
CCTTTTTTAGAGTTTGAACCCGTGGAGGATAAAAAAAAACCTCAGAAGTC 1400
CAGCGGGACGAGGCTGTTTTAAAGGTAAAGGAGCTCTTCAATGCAAAAAT
ACTCAAAGTACGAAGTAAAAGCTAAGGTCATAAAGGTGAGAATGCCCGTG 1500
GAAGAGATAGGGCTGTTTAACGCACTAATAGACGGCTTGCCCAGGTACGC
ACTCACGAGGACGAAGGAAAAGGGAAAGGGAGAAGTTTTCGTTTTAGCGA 1600
CTCCTTATAAAGTCAAGGAATTGATGGAAGCTATGGAGGGTATGAAAAAA
CACATAAAGGATTTAGAAATCCTCGGAGAGACGGATGAGGATTTAACTTT 1700
TTAAAGTATGGGTGTATCTGAGCAAAGGTTTAAGCTAAAAACAAACCTGA
AACCCGCAGGGGACCAGCCGAAAGCCATAAAAAAACTCCTTGAAAACCTA 1800
AGGAAAGGCGTAAAAGAACAAACACTTCTCGGAGTCACGGGAAGCGGAAA
GACTTTTACTCTAGCAAACGTAATAGCGAAGTACAACAAACCAACTCTTG 1900
TGGTAGTTCACAACAAAATTCTCGCGGCACAGCTATACAGGGAGTTTAAA
GAACTATTCCCTGAAAACGCTGTAGAGTACTTTGTCTCTTACTACGACTA 2000
TTACCAACCTGAAGCCTACATTCCCGAAAAAGATTTATACATAGAAAAGG
ACGCGAGTATAAACGAAAGCTGGAACGTTTCAGACACTCCGCCACGATAT 2100
CCGTTCTAGAAAGGAGGGACGTTATAGTAGTTGCTTCAGTTTCTTGCATA
TACGGACTCGGGAAACCTGAGCACTACGAAAACCTGAGGATAAAACTCCA 2200
AAGGGGAATAAGACTGAACTTGAGTAAGCTCCTGAGGAAACTCGTTGAGC
TAGGATATCAGAGAAATGACTTTGCCATAAAGAGGGCTACCTTCTCGGTT 2300
AGGGGAGACGTGGTTGAGATAGTCCCTTCTCACACGGAAGATTACCTCGT
GAGGGTAGAGTTCTGGGACGACGAAGTTGAAAGAATAGTCCTCATGGACG 2400
CTCTGAAC

FIG. 36

```
MNYVPFARKYRPKFFREVIGQEAPVRILKNAIKNDRVAHAYLFAGPRGVG
KTTIARILAKALNCKNPSKGEPCGECENCREIDRGVFPDLIEMDAASNRG    100
IDDVRALKEAVNYKPIKGKYKVYIIDEAHMLTKEAFNALLKTLEEPPPRT
VFVLCTTEYDKILPTILSRCQRIIFSKVRKEKVIEYLKKICEKEGIECEE    200
GALEVLAHASEGCMRDAASLLDQASVYGEGRVTKEVVENFLGILSQESVR
SFLKLLLNSEVDEAIKFLRELSEKGYNLTKFWEMLEEEVRNAILVKSLKN    300
PESVVQNWQDYEDFKDYPLEALLYVENLINRGKVEARTREPLRAFELAVI
KSLIVKDIIPVSQLGSVVKETKKEEKKVEVKEEPKVKEEKPKEQEEDRFQ    400
KVLNAVDGKILKRILEGAKREERDGKIVLKIEASYLRTMKKEFDSLKETF
PFLEFEPVEDKKKPQKSSGTRLF                                473
```

FIG. 37

```
ATGCGCGTTAAGGTGGACAGGGAGGAGCTTGAAGAGGTTCTTAAAAAAGC
AAGAGAAAGCACGGAAAAAAAAGCCGCACTCCCGATACTCGCGAACTTCT    100
TACTCTCCGCAAAAGAGGAAAACTTAATCGTAAGGGCAACGGACTTGGAA
AACTACCTTGTAGTCTCCGTAAAGGGGGAGGTTGAAGAGGAAGGAGAGGT    200
TTGCGTCCACTCTCAAAAACTCTACGATATAGTCAAGAACTTAAATTCCG
CTTACGTTTACCTTCATACGGAAGGTGAAAAACTCGTCATAACGGGAGGA    300
AAGAGTACGTACAAACTTCCGACAGCTCCCGCGGAGGACTTTCCCGAATT
TCCAGAAATCGTAGAAGGAGGAGAAACACTTTCGGGAAACCTTCTCGTTA    400
ACGGAATAGAAAAGGTAGAGTACGCCATAGCGAAGGAAGAAGCGAACATA
GCCCTTCAGGGAATGTATCTGAGAGGATACGAGGACAGAATTCACTTTGT    500
GTTCGGACGGTCACAGGCTTGCACTTTATGAACCTCTACGTAAACATTGA
AAAGAGTGAAGACGAGTCTTTTGCTTACTTCTCCACTCCCGAGTGGAAAC    600
TCGCCGTTAGCTCCTGGAAGGAGAATTCCCGGACTACATGAGTGTCATCC
CTGAGGAGTTTTCGGCGGAAGTCTTGTTTGAGACAGAGGAAGTCTTAAAG    700
GTTTTAAAGAGGTTGAAGGCTTTAAGCGAAGGAAAAGTTTTTCCCGTGAA
GATTACCTTAAGCGAAAACCTTGCCATCTTTGAGTTCGCGGATCCGGAGT    800
TCGGAGAAGCGAGAGAGGAAATTGAAGTGGAGTACACGGGAGAGCCCTTT
GAGATAGGATTCAACGGAAATACCTTATGGAGGCGCTTGACGCCTACGAC    900
AGCGAAAGAGTGTGGTTCAAGTTCACAACCCCCGACACGGCCACTTTATT
GGAGGCTGAAGATTACGAAAAGGAACCTTACAAGTGCATAATAATGCCGA   1000
TGAGGGTGTAGCCATGAAAAAAGCTTTAATCTTTTTATTGAGCTTGAGCC
TTTTAATTCCTGCGTTTAGCGAAGCCAAACCCAAGTCTTC             1090
```

FIG. 38

```
MRVKVDREELEEVLKKARESTEKKAALPILANFLLSAKEENLIVRATDLE
NYLVVSVKGEVEEEGEVCVHSQKLYDIVKNLNSAYVYLHTEGEKLVITGG    100
KSTYKLPTAPAEDFPEFPEIVEGGETLSGNLLVNGIEKVEYAIAKEEANI
ALQGMYLRGYEDRIHFVGSDGHRLALYEPLGEFSKELLIPRKSLKVLKKL    200
ITGIEDVNIEKSEDESFAYFSTPEWKLAVRLLEGEFPDYMSVIPEEFSAE
VLFETEEVLKVLKRLKALSEGKVFPVKITLSENLAIFEFADPEFGEAREE    300
IEVEYTGEPFEIGFNGKYLMEALDAYDSERVWFKFTTPDTATLLEAEDYE
KEPYKCIIMPMRV                                          363
```

FIG. 39

```
GTGGAAACCACAATATTCCAGTTCCAGAAAACTTTTTTCACAAAACCTCC
GAAGGAGAGGGTCTTCGTCCTTCATGGAGAAGAGCAGTATCTCATAAGAA   100
CCTTTTTGTCTAAGCTGAAGGAAAAGTACGGGGAGAATTACACGGTTCTG
TGGGGGGATGAGATAAGCGAGGAGGAATTCTACACTGCCCTTTCCGAGAC   200
CAGTATATTCGGCGGTTCAAAGGAAAAAGCGGTGGTCATTTACAACTTCG
GGGATTTCCTGAAGAAGCTCGGAAGGAAGAAAAAGGAAAAAGAAAGGCTT   300
ATAAAAGTCCTCAGAAACGTAAAGAGTAACTACGTATTTATAGTGTACGA
TGCGAAACTCCAGAAACAGGAACTTTCTTCGGAACCTCTGAAATCCGTAG   400
CGTCTTTCGGCGGTATAGTGGTAGCAAACAGGCTGAGCAAGGAGAGGATA
AAACAGCTCGTCCTTAAGAAGTTCAAAGAAAAAGGGATAAACGTAGAAAA   500
CGATGCCCTTGAATACCTTCTCCAGCTCACGGGTTACAACTTGATGGAGC
TCAAACTTGAGGTTGAAAAACTGATAGATTACGCAAGTGAAAAGAAAATT   600
TTAACACTCGATGAGGTAAAGAGAGTAGCCTTCTCAGTCTCAGAAAACGT
AAACGTATTTGAGTTCGTTGATTTACTCCTCTTAAAAGATTACGAAAAGG   700
CTCTTAAAGTTTTGGACTCCCTCATTTCCTTCGGAATACACCCCCTCCAG
ATTATGAAAATCCTGTCCTCCTATGCTCTAAAACTTTACACCCTCAAGAG   800
GCTTGAAGAGAAGGGAGAGGACCTGAATAAGGCGATGGAAAGCGTGGGAA
TAAAGAACAACTTTCTCAAGATGAAGTTCAAATCTTACTTAAAGGCAAAC   900
TCTAAAGAGGACTTGAAGAACCTAATCCTCTCCCTCCAGAGGATAGACGC
TTTTTCTAAACTTTACTTTCAGGACACAGTGCAGTTGCTGGGGATTTCTT  1000
GACCTCAAGACTGGAGAGGGAAGTTGTGAAAAATACTTCTCATGGTGGAT
AATCTTTTTTATGAAGTTTGCGGTTTGCGTTTTTCCCGGTTCT         1093
```

FIG. 40

```
VETTIFQFQKTFFTKPPKERVFVLHGEEQYLIRTFLSKLKEKYGENYTVL
WGDEISEEEFYTALSETSIFGGSKEKAVVIYNFGDFLKKLGRKKKEKERL   100
IKVLRNVKSNYVFIVYDAKLQKQELSSEPLKSVASFGGIVVANRLSKERI
KQLVLKKFKEKGINVENDALEYLLQLTGYNLMELKLEVEKLIDYASEKKI   200
LTLDEVKRVAFSVSENVNVFEFVDLLLLKDYEKALKVLDSLISFGIHPLQ
IMKILSSYALKLYTLKRLEEKGEDLNKAMESVGIKNNFLKMKFKSYLKAN   300
SKEDLKNLILSLQRIDAFSKLYFQDTVQLLRDFLTSRLEREVVKNTSHGG
```

FIG. 41

```
ATGGAAAAAGTTTTTTTGGAAAAACTCCAGAAAACCTTGCACATACCCGG
AGGACTCCTTTTTTACGGCAAAGAAGGAAGCGGAAAGACGAAAACAGCTT   100
TTGAATTTGCAAAAGGTATTTTATGTAAGGAAAACGTACCTGGGGATGCG
GAAGTTGTCCCTCCTGCAAACACGTAAACGAGCTGGAGGAAGCCTTCTTT   200
AAAGGAGAAATAGAAGACTTTAAAGTTTATAAGACAAGGACGGTAAAAAG
CACTTCGTTTACCTTATGGGCGAACATCCCGACTTTGTGGTAATAATCCC   300
GAGCGGACATTACATAAAGATAGAACAGATAAGGGAAGTTAAGAACTTTG
CCTATGTGAAGCCCGCACTAAGCAGGAGAAAAGTAATTATAATAGACGAC   400
GCCCACGCGATGACCTCTCAGGCGGCAAACGCTCTTTTAAAGGTATTGGA
AGAGCCACCTGCGGACACCACCTTTATCTTGACCACGAACAGGCGTTCTG   500
CAATCCTGCCGACTATCCTCTCCAGAACTTTTCAAGTGGAGTTCAAGGGC
TTTTCAGTAAAAGAGGTTATGGAAATAGCGAAAGTAGACGAGGAAATAGC   600
GAAACTCTCTGGAGGCAGTCTAAAAAGGGCTATCTTACTAAAGGAAAACA
AAGATATCCTAAACAAAGTAAAGGAATTCTTGGAAAACGAGCCGTTAAAA   700
GTTTACAAGCTTGCAAGTGAATTCGAAAAGTGGGAACCTGAAAAGCAAAA
ACTCTTCCTTGAAATTATGGAAGAATTGGTATCTCAAAAATTGACCGAAG   800
AGAAAAAAGACAATTACACCTACCTTCTTGATACGATCAGACTCTTTAAA
GACGGACTCGCAAGGGGTGTAAACGAACCTCTGTGGCTGTTTACGTTAGC   900
CGTTCAGGCGGATTAATAAACCGTTATTGATTCCGTAACATTTAAACCTT
AATCTAAATTATGAGAGCCTTTGAAGGAGGTCTGGTATGGAAAATTTGAA  1000
GATTAGATATATAGATACGAGGAAGATAGGAACCGTGAGCGGTGTAAAAG
T                                                    1051
```

FIG. 42

```
MEKVFLEKLQKTLHIPGGLLFYGKEGSGKTKTAFEFAKGILCKENVPWGC
GSCPSCKHVNELEEAFFKGEIEDFKVYKDKDGKKHFVYLMGEHPDFVVII   100
PSGHYIKIEQIREVKNFAYVKPALSRRKVIIIDDAHAMTSQAANALLKVL
EEPPADTTFILTTNRRSAILPTILSRTFQVEFKGFSVKEVMEIAKVDEEI   200
AKLSGGSLKRAILLKENKDILNKVKEFLENEPLKVYKLASEFEKWEPEKQ
KLFLEIMEELVSQKLTEEKKDNYTYLLDTIRLFKDGLARGVNEPLWLFTL   300
AVQAD
```

FIG. 43

```
ATGAACTTCCTGAAAAAGTTCCTTTTACTGAGAAAAGCTCAAAAGTCTCC
TTACTTCGAAGAGTTCTACGAAGAAATCGATTTGAACCAGAAGGTGAAAG   100
ATGCAAGGTTTGTAGTTTTTGACTGCGAAGCCACAGAACTCGACGTAAAG
AAGGCAAAACTCCTTTCAATAGGTGCGGTTGAGGTTAAAAACCTGGAAAT   200
AGACCTCTCTAAATCTTTTTACGAGATACTCAAAAGTGACGAGATAAAGG
CGGCGGAGATACATGGAATAACCAGGGAAGACGTTGAAAAGTACGGAAAG   300
GAACCAAAGGAAGTAATATACGACTTTCTGAAGTACATAAAGGGAAGCGT
TCTCGTTGGCTACTACGTGAAGTTTGACGTCTCACTCGTTGAGAAGTACT   400
CCATAAAGTACTTCCAGTATCCAATCATCAACTACAAGTTAGACCTGTTT
AGTTTCGTGAAGAGAGAGTACCAGAGTGGCAGGAGTCTTGACGACCTTAT   500
GAAGGAACTCGGTGTAGAAATAAGGGCAAGGCACAACGCCCTTGAAGATG
CCTACATAACCGCTCTTCTTTTCCTAAAGTACGTTTACCCGAACAGGGAG   600
TACAGACTAAAGGATCTCCCGATTTTCCTT
```

FIG. 44

```
MNFLKKFLLLRKAQKSPYFEEFYEEIDLNQKVKDARFVVFDCEATELDVK
KAKLLSIGAVEVKNLEIDLSKSFYEILKSDEIKAAEIHGITREDVEKYGK   100
EPKEVIYDFLKYIKGSVLVGYYVKFDVSLVEKYSIKYFQYPIINYKLDLF
SFVKREYQSGRSLDDLMKELGVEIRARHNALEDAYITALLFLKYVYPNRE   200
YRLKDLPIFL
```

FIG. 45

```
ATGCTCAATAAGGTTTTTATAATAGGAAGACTTACGGGTGACCCCGTTAT
AACTTATCTACCGAGCGGAACGCCCGTAGTAGAGTTTACTCTGGCTTACA   100
ACAGAAGGTATAAAAACCAGAACGGTGAATTTCAGGAGGAAAGTCACTTC
TTTGACGTAAAGGCGTACGGAAAAATGGCTGAAGACTGGGCTACACGCTT   200
CTCGAAAGGATACCTCGTACTCGTAGAGGGAAGACTCTCCCAGGAAAAGT
GGGAGAAAGAAGGAAAGAAGTTCTCAAAGGTCAGGATAATAGCGGAAAAC   300
GTAAGATTAATAAACAGGCCGAAAGGTGCTGAACTTCAAGCAGAAGAAGA
GGAGGAAGTTCCTCCCATTGAGGAGGAAATTGAAAAACTCGGTAAAGAGG   400
AAGAGAAGCCTTTTACCGATGAAGAGGACGAAATACCTTTTTAATTTTGA
GGAGGTTAAAGTATGGTAGTGAGAGCTCCTAAGAAGAAAGTTTGTATGTA   500
CTGTGAACAAAAGAGAGAGCCAGATT
```

FIG. 46

```
MLNKVFIIGRLTGDPVITYLPSGTPVVEFTLAYNRRYKNQNGEFQEESHF
FDVKAYGKMAEDWATRFSKGYLVLVEGRLSQEKWEKEGKKFSKVRIIAEN   100
VRLINRPKGAELQAEEEEEVPPIEEEIEKLGKEEEKPFTDEEDEIPF
```

FIG. 47

```
ATGCAATTTGTGGATAAACTTCCCTGTGACGAATCCGCCGAGAGGGCGGT
TCTTGGCAGTATGCTTGAAGACCCCGAAAACATACCTCTGGTACTTGAAT   100
ACCTTAAAGAAGAAGACTTCTGCATAGACGAGCACAAGCTACTTTTCAGG
GTTCTTACAAACCTCTGGTCCGAGTACGGCAATAAGCTCGATTTCGTATT   200
AATAAAGGATCACCTTGAAAAGAAAAACTTACTCCAGAAAATACCTATAG
ACTGGCTCGAAGAACTCTACGAGGAGGCGGTATCCCCTGACACGCTTGAG   300
GAAGTCTGCAAAATAGTAAAACAACGTTCCGCACAGAGGGCGATAATTCA
ACTCGGTATAGAACTCATTCACAAAGGAAAGGAAAACAAAGACTTTCACA   400
CATTAATCGAGGAAGCCCAGAGCAGGATATTTTCCATAGCGGAAAGTGCT
ACATCTACGCAGTTTTACCATGTGAAAGACGTTGCGGAAGAAGTTATAGA   500
ACTCATTTATAAATTCAAAAGCTCTGACAGGCTAGTCACGGGACTCCCAA
GCGGTTTCACGGAACTCGATCTAAAGACGACGGGATTCCACCCTGGAGAC   600
TTAATAATACTCGCCGCAAGACCCGGTATGGGGAAAACCGCCTTTATGCT
CTCCATAATCTACAATCTCGCAAAAGACGAGGGAAAACCCTCAGCTGTAT   700
TTTCCTTGGAAATGAGCAAGGAACAGCTCGTTATGAGACTCCTCTCTATG
ATGTCGGAGGTCCCACTTTTCAAGATAAGGTCTGGAAGTATATCGAATGA   800
AGATTTAAAGAAGCTTGAAGCAAGCGCAATAGAACTCGCAAAGTACGACA
TATACCTCGACGACACACCCGCTCTCACTACAACGGATTTAAGGATAAGG   900
GCAAGAAAGCTCAGAAAGGAAAAGGAAGTTGAGTTCGTGGCGGTGGACTA
CTTGCAACTTCTGAGACCGCCAGTCCGAAAGAGTTCAAGACAGGAGGAAG  1000
TGGCAGAGGTTTCAAGAAACTTAAAAGCCCTTGCAAAGGAACTTCACATT
CCCGTTATGGCACTTGCGCAGCTCTCCCGTGAGGTGGAAAAGAGGAGTGA  1100
TAAAAGACCCCAGCTTGCGGACCTCAGAGAATCCGGACAGATAGAACAGG
ACGCAGACCTAATCCTTTTCCTCCACAGACCCGAGTACTACAAGAAAAAG  1200
CCAAATCCCGAAGAGCAGGGTATAGCGGAAGTGATAATAGCCAAGCAAAG
GCAAGGACCCACGGACATTGTGAAGCTCGCATTTATTAAGGAGTACACTA  1300
AGTTTGCAAACCTAGAAGCCCTTCCTGAACAACCTCCTGAAGAAGAGGAA
CTTTCCGAAATTATTGAAACACAGGAGGATGAAGGATTCGAAGATATTGA  1400
CTTCTGAAAATTAAGGTTTTATAATTTTATCTTGGCTATCCGGGGTAGCT
CAATCGGCAGAGCGGGTGGCTG                               1472
```

FIG. 48

```
MQFVDKLPCDESAERAVLGSMLEDPENIPLVLEYLKEEDFCIDEHKLLFR
VLTNLWSEYGNKLDFVLIKDHLEKKNLLQKIPIDWLEELYEEAVSPDTLE   100
EVCKIVKQRSAQRAIIQLGITSTQFYHVKDVAEEVIELIYKFKSSDRLVT
GLPSGFTELDLKTTGFHPGDLIILAARPGMGKTAFMLSIIYNLAKDEGKP   200
SAVFSLEMSKEQLVMRLLSMMSEVPLFKIRSGSISNEDLKKLEASAIELA
KYDIYLDDTPALTTTDLRIRARKLRKEKEVEFVAVDYLQLLRPPVRKSSR   300
QEEVAEVSRNLKALAKELHIPVMALAQLSREVEKRSDKRPQLADLRESGQ
IEQDADLILFLHRPEYYKKKPNPEEQGIAEVIIAKQRQGPTDIVKLAFIK   400
EYTKFANLEALPEQPPEEEELSEIIETQEDEGFEDIDF
```

FIG. 49

ATGTCCTCGGACATAGACGAACTTAGACGGGAAATAGATATAGTAGACGT
CATTTCCGAATACTTAAACTTAGAGAAGGTAGGTTCCAATTACAGAACGA 100
ACTGTCCCTTTCACCCTGACGATACACCCTCCTTTTACGTGTCTCCAAGT
AAACAAATATTCAAGTGTTTCGGTTGCGGGGTAGGGGGAGACGCGATAAA 200
GTTCGTTTCCCTTTACGAGGACATCTCCTATTTTGAAGCCGCCCTTGAAC
TCGCAAAACGCTACGGAAAGAAATTAGACCTTGAAAAGATATCAAAAGAC 300
GAAAAGGTATACGTGGCTCTTGACAGGGTTTGTGATTTCTACAGGGAAAG
CCTTCTCAAAAACAGAGAGGCAAGTGAGTACGTAAAGAGTAGGGGAATAG 400
ACCCTAAAGTAGCGAGGAAGTTTGATCTTGGGTACGCACCTTCCAGTGAA
GCACTCGTAAAAGTCTTAAAAGAGAACGATCTTTTAGAGGCTTACCTTGA 500
AACTAAAAACCTCCTTTCTCCTACGAAGGGTGTTTACAGGGATCTCTTTC
TTCGGCGTGTCGTGATCCCGATAAAGGATCCGAGGGGAAGAGTTATAGGT 600
TTCGGTGGAAGGAGGATAGTAGAGGACAAATCTCCCAAGTACATAAACTC
TCCAGACAGCAGGGTATTTAAAAAGGGGGAGAACTTATTCGGTCTTTACG 700
AGGCAAAGGAGTATATAAAGGAAGAAGGATTTGCGATACTTGTGGAAGGG
TACTTTGACCTTTTGAGACTTTTTTCCGAGGGAATAAGGAACGTTGTTGC 800
ACCCCTCGGTACAGCCCTGACCCAAAATCAGGCAAACCTCCTTTCCAAGT
TCACAAAAAAGGTCTACATCCTTTACGACGGAGATGATGCGGGAAGAAAG 900
GCTATGAAAAGTGCCATTCCCCTACTCCTCAGTGCAGGAGTGGAAGTTTA
TCCCGTTTACCTCCCCGAAGGATACGATCCCGACGAGTTTATAAAGGAAT 1000
TCGGGAAAGAGGAATTAAGAAGACTGATAAACAGCTCAGGGGAGCTCTTT
GAAACGCTCATAAAAACCGCAAGGGAAAACTTAGAGGAGAAAACGCGTGA 1100
GTTCAGGTATTATCTGGGCTTTATTTCCGATGGAGTAAGGCGCTTTGCTC
TGGCTTCGGAGTTTCACACCAAGTACAAAGTTCCTATGGAAATTTTATTA 1200
ATGAAAATTGAAAAAAATTCTCAAGAAAAAGAAATTAAACTCTCCTTTAA
GGAAAAAATCTTCCTGAAAGGACTGATAGAATTAAAACCAAAAATAGACC 1300
TTGAAGTCCTGAACTTAAGTCCTGAGTTAAAGGAACTCGCAGTTAACGCC
TTAAACGGAGAGGAGCATTTACTTCCAAAAGAAGTTCTCGAGTACCAGGT 1400
GGATAACTTGGAGAAACTTTTTAACAACATCCTTAGGGATTTACAAAAAT
CTGGGAAAAAGAGGAAGAAAAGAGGGTTGAAAAATGTAAATACTTAATTA 1500
ACTTTAATAAATTTTTAGAGTTAGGA

FIG. 50

MSSDIDELRREIDIVDVISEYLNLEKVGSNYRTNCPFHPDDTPSFYVSPS
KQIFKCFGCGVGGDAIKFVSLYEDISYFEAALELAKRYGKKLDLEKISKD 100
EKVYVALDRVCDFYRESLLKNREASEYVKSRGIDPKVARKFDLGYAPSSE
ALVKVLKENDLLEAYLETKNLLSPTKGVYRDLFLRRVVIPIKDPRGRVIG 200
FGGRRIVEDKSPKYINSPDSRVFKKGENLFGLYEAKEYIKEEGFAILVEG
YFDLLRLFSEGIRNVVAPLGTALTQNQANLLSKFTKKVYILYDGDDAGRK 300
AMKSAIPLLLSAGVEVYPVYLPEGYDPDEFIKEFGKEELRRLINSSGELF
ETLIKTARENLEEKTREFRYYLGFISDGVRRFALASEFHTKYKVPMEILL 400
MKIEKNSQEKEIKLSFKEKIFLKGLIELKPKIDLEVLNLSPELKELAVNA
LNGEEHLLPKEVLEYQVDNLEKLFNNILRDLQKSGKKRKKRGLKNVNT 498

FIG. 51

```
ATGCAAGATACCGCTACCTGCAGTATTTGTCAGGGGACGGGATTCGTAAA
GACCGAAGACAACAAGGTAAGGCTCTGCGAATGCAGGTTCAAGAAAAGGG    100
ATGTAAACAGGGAACTAAACATCCCAAAGAGGTACTGGAACGCCAACTTA
GACACTTACCACCCCAAGAACGTATCCCAGAACAGGGCACTTTTGACGAT    200
AAGGGTCTTCGTCCACAACTTCAATCCCGAGGAAGGGAAAGGGCTTACCT
TTGTAGGATCTCCTGGAGTCGGCAAAACTCACCTTGCGGTTGCAACATTA    300
AAAGCGATTTATGAGAAGAAGGGAATCAGAGGATACTTCTTCGATACGAA
GGATCTAATATTCAGGTTAAAACACTTAATGGACGAGGGAAAGGATACAA    400
AGTTTTTAAAAACTGTCTTAAACTCACCGGTTTTGGTTCTCGACGACCTC
GGTTCTGAGAGGCTCAGTGACTGGCAGAGGGAACTCATCTCTTACATAAT    500
CACTTACAGGTATAACAACCTTAAGAGCACGATAATAACCACGAATTACT
CACTCCAGAGGGAAGAAGAGAGTAGCGTGAGGATAAGTGCGGATCTTGCA    600
AGCAGACTCGGAGAAAACGTAGTTTCAAAAATTTACGAGATGAACGAGTT
GCTCGTTATAAAGGGTTCCGACCTCAGGAAGTCTAAAAAGCTATCAACCC    700
CATCT
```

FIG. 52

```
MQDTATCSICQGTGFVKTEDNKVRLCECRFKKRDVNRELNIPKRYWNANL
DTYHPKNVSQNRALLTIRVFVHNFNPEEGKGLTFVGSPGVGKTHLAVATL    100
KAIYEKKGIRGYFFDTKDLIFRLKHLMDEGKDTKFLKTVLNSPVLVLDDL
GSERLSDWQRELISYIITYRYNNLKSTIITTNYSLQREEESSVRISADLA    200
SRLGENVVSKIYEMNELLVIKGSDLRKSKKLSTPS
```

FIG. 53

ATGAAAAAGATTGAAAATTTGAAGTGGAAAAATGTCTCGTTTAAAAGCCT
GGAAATAGATCCCGATGCAGGTGTGGTTCTCGTTTCCGTGGAAAAATTCT 100
CCGAAGAGATAGAAGACCTTGTGCGTTTACTGGAGAAGAAGACGCGGTTT
CGAGTCATCGTGAACGGTGTTCAAAAAAGTAACGGGGATCTAAGGGGAAA 200
GATACTTTCCCTTCTCAACGGTAATGTGCCTTACATAAAAGATGTTGTTT
TCGAAGGAAACAGGCTGATTCTGAAAGTGCTTGGAGATTTCGCGCGGGAC 300
AGGATCGCCTCCAAACTCAGAAGCACGAAAAAACAGCTCGATGAACTGCT
GCCTCCCGGAACAGAGATCATGCTGGAGGTTGTGGAGCCTCCGGAAGATC 400
TTTTGAAAAAGGAAGTACCACAACCAGAAAAGAGAGAAGAACCAAAGGGT
GAAGAATTGAAGATCGAGGATGAAAACCACATCTTTGGACAGAAACCCAG 500
AAAGATCGTCTTCACCCCCTCAAAAATCTTTGAGTACAACAAAAAGACAT
CGGTGAAGGGCAAGATCTTCAAAATAGAGAAGATCGAGGGGAAAAGAACG 600
GTCCTTCTGATTTACCTGACAGACGGAGAAGATTCTCTGATCTGCAAAGT
CTTCAACGACGTTGAAAAGGTCGAAGGGAAAGTATCGGTGGGAGACGTGA 700
TCGTTGCCACAGGAGACCTCCTTCTCGAAAACGGGGAGCCCACCCTTTAC
GTGAAGGGAATCACAAAACTTCCCGAAGCGAAAAGGATGGACAAATCTCC 800
GGTTAAGAGGGTGGAGCTCCACGCCCATACCAAGTTCAGCGATCAGGACG
CAATAACAGATGTGAACGAATATGTGAAACGAGCCAAGGAATGGGGCTTT 900
CCCGCGATAGCCCTCACGGATCATGGGAACGTTCAGGCCATACCTTACTT
CTACGACGCGGCGAAAGAAGCTGGAATAAAGCCCATTTTCGGTATCGAAG 1000
CGTATCTGGTGAGTGACGTGGAGCCCGTCATAAGGAATCTCTCCGACGAT
TCGACGTTTGGAGATGCCACGTTCGTCGTCCTCGACTTCGAGACGACGGG 1100
TCTCGACCCGCAGGTGGATGAGATCATCGAGATAGGAGCGGTGAAGATAC
AGGGTGGCCAGATAGTGGACGAGTACCACACTCTCATAAAGCCTTCCAGG 1200
GAGATCTCAAGAAAAAGTTCGGAGATCACCGGAATCACTCAAGAGATGCT
GGAAAACAAGAGAAGCATCGAGGAAGTTCTGCCGGAGTTCCTCGGTTTTC 1300
TGGAAGATTCCATCATCGTAGCACACAACGCCAACTTCGACTACAGATTT
CTGAGGCTGTGGATCAAAAAAGTGATGGGATTGGACTGGGAAAGACCCTA 1400
CATAGATACGCTCGCCCTCGCAAAGTCCCTTCTCAAACTGAGAAGCTACT
CTCTGGATTCCGTTGTGGAAAAGCTCGGATTGGGTCCCTTCCGGCACCAC 1500
AGGGCCCTGGATGACGCGAGGGTCACCGCTCAGGTTTTCCTCAGGTTCGT
TGAGATGATGAAGAAGATCGGTATCACGAAGCTTTCAGAAATGGAGAAGT 1600
TGAAGGATACGATAGACTACACCGCGTTGAAACCCTTCCACTGCACGATC
CTCGTTCAGAACAAAAAGGGATTGAAAAACCTATACAAACTGGTTTCTGA 1700
TTCCTATATAAAGTACTTCTACGGTGTTCCGAGGATCCTCAAAAGTGAGC
TCATCGAGAACAGAGAAGGACTGCTCGTGGGTAGCGCGTGTATCTCCGGT 1800
GAGCTCGGACGTGCCGCCCTCGAAGGAGCGAGTGATTCAGAACTCGAAGA
GATCGCGAAGTTCTACGACTACATAGAAGTCATGCCGCTCGACGTTATAG 1900
CCGAAGATGAAGAAGACCTAGACAGAGAAAGACTGAAAGAAGTGTACCGA
AAACTCTACAGAATAGCGAAAAAATTGAACAAGTTCGTCGTCATGACCGG 2000
TGATGTTCATTTCCTCGATCCCGAAGATGCCAGGGGCAGAGCTGCACTTC
TGGCACCTCAGGGAAACAGAAACTTCGAGAATCAGCCCGCACTCTACCTC 2100
AGAACGACCGAAGAAATGCTCGAGAAGGCGATAGAGATATTCGAAGATGA
AGAGATCGCGAGGGAAGTCGTGATAGAGAATCCCAACAGAATAGCCGATA 2200
TGATCGAGGAAGTGCAGCCGCTCGAGAAAAAACTTCACCCGCCGATCATA
GAGAACGCCGATGAAATAGTGAGAAACCTCACCATGAAGCGGGCGTACGA 2300
GATCTACGGTGATCCGCTTCCCGAAATCGTCCAGAAGCGTGTGGAAAAGG

FIG. 54A

AACTGAACGCCATCATAAATCATGGATACGCCGTTCTCTATCTCATCGCT 2400
CAGGAGCTCGTTCAGAAATCTATGAGCGATGGTTACGTGGTTGGATCCAG
AGGATCCGTCGGGTCTTCACTCGTGGCCAATCTCCTCGGAATAACAGAGG 2500
TGAATCCCCTACCACCACATTACAGGTGTCCAGAGTGCAAATACTTTGAA
GTTGTCGAAGACGACAGATACGGAGCGGGTTACGACCTTCCCAACAAGAA 2600
CTGTCCAAGATGTGGGGCTCCTCTCAGAAAAGACGGCCACGGCATACCGT
TTGAAACGTTCATGGGGTTCGAGGGTGACAAGGTCCCCGACATAGATCTC 2700
AACTTCTCAGGAGAGTATCAGGAACGTGCTCATCGTTTTGTGGAAGAACT
CTTCGGTAAAGACCACGTCTATAGGGCGGGAACCATAAACACCATCGCGG 2800
AAAGAAGTGCGGTGGGTTACGTGAGAAGCTACGAAGAGAAAACCGGAAAG
AAGCTCAGAAAGGCGGAAATGGAAAGACTCGTTTCCATGATCACGGGAGT 2900
GAAGAGAACGACGGGTCAGCACCCAGGGGGGCTCATGATCATACCGAAAG
ACAAAGAAGTCTACGATTTCACTCCCATACAGTATCCAGCCAACGATAGA 3000
AACGCAGGTGTGTTCACCACGCACTTCGCATACGAGACGATCCATGATGA
CCTGGTGAAGATAGATGCGCTCGGCCACGATGATCCCACTTTCATCAAGA 3100
TGCTCAAGGACCTCACCGGAATCGATCCCATGACGATTCCCATGGATGAC
CCCGATACGCTCGCCATATTCAGTTCTGTGAAGCCTCTTGGTGTGGATCC 3200
CGTTGAGCTGGAAAGCGATGTGGGAACGTACGGAATTCCGGAGTTCGGAA
CCGAGTTTGTGAGGGGAATGCTCGTTGAAACGAGACCAAAGAGTTTCGCC 3300
GAGCTTGTGAGAATCTCAGGACTGTCACACGGTACGGACGTCTGGTTGAA
CAACGCACGTGATTGGATAAACCTCGGCTACGCCAAGCTCTCCGAGGTTA 3400
TCTCGTGTAGGGACGACATCATGAACTTCCTCATACACAAAGGAATGGAA
CCGTCACTTGCCTTCAAGATCATGGAAAACGTCAGGAAGGGAAAGGGTAT 3500
CACAGAAGAGATGGAGAGCGAGATGAGAAGGCTGAAGGTTCCAGAATGGT
TCATCGAATCCTGTAAAAGGATCAAATATCTCTTCCCGAAAGCTCACGCT 3600
GTGGCTTACGTGAGTATGGCCTTCAGAATTGCTTACTTCAAGGTTCACTA
TCCTCTTCAGTTTTACGCGGCGTACTTCACGATAAAAGGTGATCAGTTCG 3700
ATCCGGTTCTCGTACTCAGGGGAAAAGAAGCCATAAAGAGGCGCTTGAGA
GAACTCAAAGCGATGCCTGCCAAAGACGCCCAGAAGAAAAACGAAGTGAG 3800
TGTTCTGGAGGTTGCCCTGGAAATGATACTGAGAGGTTTTTCCTTCCTAC
CGCCCGACATCTTCAAATCCGACGCGAAGAAATTTCTGATAGAAGGAAAC 3900
TCGCTGAGAATTCCGTTCAACAAACTTCCAGGACTGGGTGACAGCGTTGC
CGAGTCGATAATCAGAGCCAGGGAAGAAAAGCCGTTCACTTCGGTGGAAG 4000
ATCTCATGAAGAGGACCAAGGTCAACAAAAATCACATAGAGCTGATGAAA
AGCCTGGGTGTTCTCGGGGACCTTCCAGAGACGGAACAGTTCACGCTTTT 4100
C

FIG. 54B

```
MKKIENLKWKNVSFKSLEIDPDAGVVLVSVEKFSEEIEDLVRLLEKKTRF
RVIVNGVQKSNGDLRGKILSLLNGNVPYIKDVVFEGNRLILKVLGDFARD   100
RIASKLRSTKKQLDELLPPGTEIMLEVVEPPEDLLKKEVPQPEKREEPKG
EELKIEDENHIFGQKPRKIVFTPSKIFEYNKKTSVKGKIFKIEKIEGKRT   200
VLLIYLTDGEDSLICKVFNDVEKVEGKVSVGDVIVATGDLLLENGEPTLY
VKGITKLPEAKRMDKSPVKRVELHAHTKFSDQDAITDVNEYVKRAKEWGF   300
PAIALTDHGNVQAIPYFYDAAKEAGIKPIFGIEAYLVSDVEPVIRNLSDD
STFGDATFVVLDFETTGLDPQVDEIIEIGAVKIQGGQIVDEYHTLIKPSR   400
EISRKSSEITGITQEMLENKRSIEEVLPEFLGFLEDSIIVAHNANFDYRF
LRLWIKKVMGLDWERPYIDTLALAKSLLKLRSYSLDSVVEKLGLGPFRHH   500
RALDDARVTAQVFLRFVEMMKKIGITKLSEMEKLKDTIDYTALKPFHCTI
LVQNKKGLKNLYKLVSDSYIKYFYGVPRILKSELIENREGLLVGSACISG   600
ELGRAALEGASDSELEEIAKFYDYIEVMPLDVIAEDEEDLDRERLKEVYR
KLYRIAKKLNKFVVMTGDVHFLDPEDARGRAALLAPQGNRNFENQPALYL   700
RTTEEMLEKAIEIFEDEEIAREVVIENPNRIADMIEEVQPLEKKLHPPII
ENADEIVRNLTMKRAYEIYGDPLPEIVQKRVEKELNAIINHGYAVLYLIA   800
QELVQKSMSDGYVVGSRGSVGSSLVANLLGITEVNPLPPHYRCPECKYFE
VVEDDRYGAGYDLPNKNCPRCGAPLRKDGHGIPFETFMGFEGDKVPDIDL   900
NFSGEYQERAHRFVEELFGKDHVYRAGTINTIAERSAVGYVRSYEEKTGK
KLRKAEMERLVSMITGVKRTTGQHPGGLMIIPKDKEVYDFTPIQYPANDR  1000
NAGVFTTHFAYETIHDDLVKIDALGHDDPTFIKMLKDLTGIDPMTIPMDD
PDTLAIFSSVKPLGVDPVELESDVGTYGIPEFGTEFVRGMLVETRPKSFA  1100
ELVRISGLSHGTDVWLNNARDWINLGYAKLSEVISCRDDIMNFLIHKGME
PSLAFKIMENVRKGKGITEEMESEMRRLKVPEWFIESCKRIKYLFPKAHA  1200
VAYVSMAFRIAYFKVHYPLQFYAAYFTIKGDQFDPVLVLRGKEAIKRRLR
ELKAMPAKDAQKKNEVSVLEVALEMILRGFSFLPPDIFKSDAKKFLIEGN  1300
SLRIPFNKLPGLGDSVAESIIRAREEKPFTSVEDLMKRTKVNKNHIELMK
SLGVLGDLPETEQFTLF                                    1367
```

FIG. 55

```
GTGCTCGCCATGATATGGAACGACACCGTTTTTTGCGTCGTAGACACAGA
AACCACGGGAACCGATCCCTTTGCCGGAGACCGGATAGTTGAAATAGCCG    100
CTGTTCCTGTCTTCAAGGGGAAGATCTACAGAAACAAAGCGTTTCACTCT
CTCGTGAATCCCAGAATAAGAATCCCTGCGCTGATTCAGAAAGTTCACGG    200
TATCAGCAACATGGACATCGTGGAAGCGCCAGACATGGACACAGTTTACG
ATCTTTTCAGGGATTACGTGAAGGGAACGGTGCTCGTGTTTCACAACGCC    300
AACTTCGACCTCACTTTTCTGGATATGATGGCAAAGGAAACGGGAAACTT
TCCAATAACGAATCCCTACATCGACACACTCGATCTTTCAGAAGAGATCT    400
TTGGAAGGCCTCATTCTCTCAAATGGCTCTCCGAAAGACTTGGAATAAAA
ACCACGATACGGCACCGTGCTCTTCCAGATGCCCTGGTGACCGCAAGAGT    500
TTTTGTGAAGCTTGTTGAATTTCTTGGTGAAAACAGGGTCAACGAATTCA
TACGTGGAAAACGGGGG                                     567
```

FIG. 56

```
MLAMIWNDTVFCVVDTETTGTDPFAGDRIVEIAAVPVFKGKIYRNKAFHS
LVNPRIRIPALIQKVHGISNMDIVEAPDMDTVYDLFRDYVKGTVLVFHNA    100
NFDLTFLDMMAKETGNFPITNPYIDTLDLSEEIFGRPHSLKWLSERLGIK
TTIRHRALPDALVTARVFVKLVEFLGENRVNEFIRGKRG               189
```

FIG. 57

GTGGAAGTTCTTTACAGGAAGTACAGGCCAAAGACTTTTTCTGAGGTTGT
CAATCAGGATCATGTGAAGAAGGCAATAATCGGTGCTATTCAGAAGAACA 100
GCGTGGCCCACGGATACATATTCGCCGGTCCGAGGGGAACGGGGAAGACT
ACTCTTGCCAGAATTCTCGCAAAATCCCTGAACTGTGAGAACAGAAAGGG 200
AGTTGAACCCTGCAATTCCTGCAGAGCCTGCAGAGAGATAGACGAGGGAA
CCTTCATGGACGTGATAGAGCTCGACGCGGCCTCCAACAGAGGAATAGAC 300
GAGATCAGAAGAATCAGAGACGCCGTTGGATACAGGCCGATGGAAGGTAA
ATACAAAGTCTACATAATAGACGAAGTTCACATGCTCACGAAAGAAGCCT 400
TCAACGCGCTCCTCAAAACACTCGAAGAACCTCCTTCCCACGTCGTGTTC
GTGCTGGCAACGACAAACCTTGAGAAGGTTCCTCCCACGATTATCTCGAG 500
ATGTCAGGTTTTCGAGTTCAGAAACATTCCCGACGAGCTCATCGAAAAGA
GGCTCCAGGAAGTTGCGGAGGCTGAAGGAATAGAGATAGACAGGGAAGCT 600
CTGAGCTTCATCGCAAAAAGAGCCTCTGGAGGCTTGAGAGACGCGCTCAC
CATGCTCGAGCAGGTGTGGAAGTTCTCGGAAGGAAAGATAGATCTCGAGA 700
CGGTACACAGGGCGCTCGGGTTGATACCGATACAGGTTGTTCGCGATTAC
GTGAACGCTATCTTTTCTGGTGATGTGAAAAGGGTCTTCACCGTTCTCGA 800
CGACGTCTATTACAGCGGGAAGGACTACGAGGTGCTCATTCAGGAAGCAG
TCGAGGATCTGGTCGAAGACCTGGAAAGGGAGAGAGGGGTTTACCAGGTT 900
TCAGCGAACGATATAGTTCAGGTTTCGAGACAACTTCTGAATCTTCTGAG
AGAGATAAAGTTCGCCGAAGAAAAACGACTCGTCTGTAAAGTGGGTTCGG 1000
CTTACATAGCGACGAGGTTCTCCACCACAAACGTTCAGGAAAACGATGTC
AGAGAAAAAAACGATAATTCAAATGTACAGCAGAAAGAAGAGAAGAAAGA 1100
AACGGTGAAGGCAAAAGAAGAAAAACAGGAAGACAGCGAGTTCGAGAAAC
GCTTCAAAGAACTCATGGAAGAACTGAAAGAAAAGGGCGATCTCTCTATC 1200
TTTGTCGCTCTCAGCCTCTCAGAGGTGCAGTTTGACGGAGAAAAGGTGAT
TATTTCTTTTGATTCATCGAAAGCTATGCATTACGAGTTGATGAAGAAAA 1300
AACTGCCTGAGCTGGAAAACATTTTTTCTAGAAAACTCGGGAAAAAAGTA
GAAGTTGAACTTCGACTGATGGGAAAAGAAGAAACAATCGAGAAGGTTTC 1400
TCAGAAGATCCTGAGATTGTTTGAACAGGAGGGA

FIG. 58

MEVLYRKYRPKTFSEVVNQDHVKKAIIGAIQKNSVAHGYIFAGPRGTGKT
TLARILAKSLNCENRKGVEPCNSCRACREIDEGTFMDVIELDAASNRGID 100
EIRRIRDAVGYRPMEGKYKVYIIDEVHMLTKEAFNALLKTLEEPPSHVVF
VLATTNLEKVPPTIISRCQVFEFRNIPDELIEKRLQEVAEAEGIEIDREA 200
LSFIAKRASGGLRDALTMLEQVWKFSEGKIDLETVHRALGLIPIQVVRDY
VNAIFSGDVKRVFTVLDDVYYSGKDYEVLIQEAVEDLVEDLERERGVYQV 300
SANDIVQVSRQLLNLLREIKFAEEKRLVCKVGSAYIATRFSTTNVQENDV
REKNDNSNVQQKEEKKETVKAKEEKQEDSEFEKRFKELMEELKEKGDLSI 400
FVALSLSEVQFDGEKVIISFDSSKAMHYELMKKKLPELENIFSRKLGKKV
EVELRLMGKEETIEKVSQKILRLFEQEG 478

FIG. 59

```
ATGAAAGTAACCGTCACGACTCTTGAATTGAAAGACAAAATAACCATCGC
CTCAAAAGCGCTCGCAAAGAAATCCGTGAAACCCATTCTTGCTGGATTTC   100
TTTTCGAAGTGAAAGATGGAAATTTCTACATCTGCGCGACCGATCTCGAG
ACCGGAGTCAAAGCAACCGTGAATGCCGCTGAAATCTCCGGTGAGGCACG   200
TTTTGTGGTACCAGGAGATGTCATTCAGAAGATGGTCAAGGTTCTCCCAG
ATGAGATAACGGAACTTTCTTTAGAGGGGGATGCTCTTGTTATAAGTTCT   300
GGAAGCACCGTTTTCAGGATCACCACCATGCCCGCGGACGAATTTCCAGA
GATAACGCCTGCCGAGTCTGGAATAACCTTCGAAGTTGACACTTCGCTCC   400
TCGAGGAAATGGTTGAAAAGGTCATCTTCGCCGCTGCCAAAGACGAGTTC
ATGCGAAATCTGAATGGAGTTTTCTGGGAACTCCACAAGAATCTTCTCAG   500
GCTGGTTGCAAGTGATGGTTTCAGACTTGCACTTGCTGAAGAGCAGATAG
AAAACGAGGAAGAGGCGAGTTTCTTGCTCTCTTTGAAGAGCATGAAAGAA   600
GTTCAAAACGTGCTGGACAACACAACGGAGCCGACTATAACGGTGAGGTA
CGATGGAAGAAGGGTTTCTCTGTCGACAAATGATGTAGAAACGGTGATGA   700
GAGTGGTCGACGCTGAATTTCCCGATTACAAAAGGGTGATCCCCGAAACT
TTCAAAACGAAAGTGGTGGTTTCCAGAAAAGAACTCAGGGAATCTTTGAA   800
GAGGGTGATGGTGATTGCCAGCAAGGGAAGCGAGTCCGTGAAGTTCGAAA
TAGAAGAAAACGTTATGAGACTTGTGAGCAAGAGCCCGGATTATGGAGAA   900
GTGGTCGATGAAGTTGAAGTTCAAAAAGAAGGGGAAGATCTCGTGATCGC
TTTCAACCCGAAGTTCATCGAGGACGTTTTGAAGCACATTGAGACTGAAG  1000
AAATCGAAATGAACTTCGTTGATTCTACCAGTCCATGTCAGATAAATCCA
CTCGATATTTCTGGATACCTTTACATAGTGATGCCCATCAGACTGGCA    1098
```

FIG. 60

```
MKVTVTTLELKDKITIASKALAKKSVKPILAGFLFEVKDGNFYICATDLE
TGVKATVNAAEISGEARFVVPGDVIQKMVKVLPDEITELSLEGDALVISS   100
GSTVFRITTMPADEFPEITPAESGITFEVDTSLLEEMVEKVIFAAAKDEF
MRNLNGVFWELHKNLLRLVASDGFRLALAEEQIENEEEASFLLSLKSMKE   200
VQNVLDNTTEPTITVRYDGRRVSLSTNDVETVMRVVDAEFPDYKRVIPET
FKTKVVVSRKELRESLKRVMVIASKGSESVKFEIEENVMRLVSKSPDYGE   300
VVDEVEVQKEGEDLVIAFNPKFIEDVLKHIETEEIEMNFVDSTSPCQINP
LDISGYLYIVMPIRLA                                     366
```

FIG. 61

```
ATGCCAGTCACGTTTCTCACAGGTACTGCAGAAACTCAGAAGGAAGAATT
GATAAAGAAACTCCTGAAGGATGGTAACGTGGAGTACATAAGGATCCATC    100
CGGAGGATCCCGACAAGATCGATTTCATAAGGTCTTTACTCAGGACAAAG
ACGATCTTTTCCAACAAGACGATCATTGACATCGTCAATTTCGATGAGTG    200
GAAAGCACAGGAGCAGAAGCGTCTCGTTGAACTTTTGAAAAACGTACCGG
AAGACGTTCATATCTTCATCCGTTCTCAAAAAACAGGTGGAAAGGGAGTA    300
GCGCTGGAGCTTCCGAAGCCATGGGAAACGGACAAGTGGCTTGAGTGGAT
AGAAAAGCGCTTCAGGGAGAATGGTTTGCTCATCGATAAAGATGCCCTTC    400
AGCTGTTTTTCTCCAAGGTTGGAACGAACGACCTGATCATAGAAAGGGAG
ATTGAAAAACTGAAAGCTTATTCCGAGGACAGAAAGATAACGGTAGAAGA    500
CGTGGAAGAGGTCGTTTTTACCTATCAGACTCCGGGATACGATGATTTTT
GCTTTGCTGTTTCCGAAGGAAAAAGGAAGCTCGCTCACTCTCTTCTGTCG    600
CAGCTGTGGAAAACCACAGAGTCCGTGGTGATTGCCACTGTCCTTGCGAA
TCACTTCTTGGATCTCTTCAAAATCCTCGTTCTTGTGACAAAGAAAAGAT    700
ACTACACCTGGCCTGATGTGTCCAGGGTGTCCAAAGAGCTGGGAATTCCC
GTTCCTCGTGTGGCTCGTTTCCTCGGTTTCTCCTTTAAGACCTGGAAATT    800
CAAGGTGATGAACCACCTCCTCTACTACGATGTGAAGAAGGTTAGAAAGA
TACTGAGGGATCTCTACGATCTGGACAGAGCCGTGAAAAGCGAAGAAGAT    900
CCAAAACCGTTCTTCCACGAGTTCATAGAAGAGGTGGCACTGGATGTATA
TTCTCTTCAGAGAGATGAAGAA                                 972
```

FIG. 62

```
MPVTFLTGTAETQKEELIKKLLKDGNVEYIRIHPEDPDKIDFIRSLLRTK
TIFSNKTIIDIVNFDEWKAQEQKRLVELLKNVPEDVHIFIRSQKTGGKGV    100
ALELPKPWETDKWLEWIEKRFRENGLLIDKDALQLFFSKVGTNDLIIERE
IEKLKAYSEDRKITVEDVEEVVFTYQTPGYDDFCFAVSEGKRKLAHSLLS    200
QLWKTTESVVIATVLANHFLDLFKILVLVTKKRYYTWPDVSRVSKELGIP
VPRVARFLGFSFKTWKFKVMNHLLYYDVKKVRKILRDLYDLDRAVKSEED    300
PKPFFHEFIEEVALDVYSLQRDEE
```

FIG. 63

ATGAACGATTTGATCAGAAAGTACGCTAAAGATCAACTGGAAACTTTGAA
AAGGATCATAGAAAAGTCTGAAGGAATATCCATCCTCATAAATGGAGAAG 100
ATCTCTCGTATCCGAGAGAAGTATCCCTTGAACTTCCCGAGTACGTGGAG
AAATTTCCCCCGAAGGCCTCGGATGTTCTGGAGATAGATCCCGAGGGGGA 200
GAACATAGGCATAGACGACATCAGAACGATAAAGGACTTCCTGAACTACA
GCCCCGAGCTCTACACGAGAAAGTACGTGATAGTCCACGACTGTGAAAGA 300
ATGACCCAGCAGGCGGCGAACGCGTTTCTGAAGGCCCTTGAAGAACCACC
AGAATACGCTGTGATCGTTCTGAACACTCGCCGCTGGCATTATCTACTGC 400
CGACGATAAAGAGCCGAGTGTTCAGAGTGGTTGTGAACGTTCCAAAGGAG
TTCAGAGATCTCGTGAAAGAGAAAATAGGAGATCTCTGGGAGGAACTTCC 500
ACTTCTTGAGAGAGACTTCAAAACGGCTCTCGAAGCCTACAAACTTGGTG
CGGAAAAACTTTCTGGATTGATGGAAAGTCTCAAAGTTTTGGAGACGGAA 600
AAACTCTTGAAAAAGGTCCTTTCAAAAGGCCTCGAAGGTTATCTCGCATG
TAGGGAGCTCCTGGAGAGATTTTCAAAGGTGGAATCGAAGGAATTCTTTG 700
CGCTTTTTGATCAGGTGACTAACACGATAACAGGAAAAGACGCGTTTCTT
TTGATCCAGAGACTGACAAGAATCATTCTCCACGAAAACACATGGGAAAG 800
CGTTGAAGATCAAAAAAGCGTGTCTTTCCTCGATTCAATTCTCAGGGTGA
AGATAGCGAATCTGAACAACAAACTCACTCTGATGAACATCCTCGCGATA 900
CACAGAGAGAGAAAGAGAGGTGTCAACGCTTGGAGC

FIG. 64

MNDLIRKYAKDQLETLKRIIEKSEGISILINGEDLSYPREVSLELPEYVE
KFPPKASDVLEIDPEGENIGIDDIRTIKDFLNYSPELYTRKYVIVHDCER 100
MTQQAANAFLKALEEPPEYAVIVLNTRRWHYLLPTIKSRVFRVVVNVPKE
FRDLVKEKIGDLWEELPLLERDFKTALEAYKLGAEKLSGLMESLKVLETE 200
KLLKKVLSKGLEGYLACRELLERFSKVESKEFFALFDQVTNTITGKDAFL
LIQRLTRIILHENTWESVEDKSVSFLDSILRVKIANLNNKLTLMNILAIH 300
RERKRGVNAWS

FIG. 65

```
ATGTCTTTCTTCAACAAGATCATACTCATAGGAAGACTCGTGAGAGATCC
CGAAGAGAGATACACGCTCAGCGGAACTCCAGTCACCACCTTCACCATAG    100
CGGTGGACAGGGTTCCCAGAAAGAACGCGCCGGACGACGCTCAAACGACT
GATTTCTTCAGGATCGTCACCTTTGGAAGACTGGCAGAGTTCGCTAGAAC    200
CTATCTCACCAAAGGAAGGCTCGTTCTCGTCGAAGGTGAAATGAGAATGA
GAAGATGGGAAACACCCACTGGAGAAAAGAGGGTATCTCCGGAGGTTGTC    300
GCAAACGTTGTTAGATTCATGGACAGAAAACCTGCTGAAACAGTTAGCGA
GACTGAAGAGGAGCTGGAAATACCGGAAGAAGACTTTTCCAGCGATACCT    400
TCAGTGAAGATGAACCACCATTT
```

FIG. 66

```
MSFFNKIILIGRLVRDPEERYTLSGTPVTTFTIAVDRVPRKNAPDDAQTT
DFFRIVTFGRLAEFARTYLTKGRLVLVEGEMRMRRWETPTGEKRVSPEVV    100
ANVVRFMDRKPAETVSETEEELEIPEEDFSSDTFSEDEPPF
```

FIG. 67

```
ATGCGTGTTCCCCCGCACAACTTAGAGGCCGAAGTTGCTGTGCTCGGAAG
CATATTGATAGATCCGTCGGTAATAAACGACGTTCTTGAAATTTTGAGCC      100
ACGAAGATTTCTATCTGAAAAAACACCAACACATCTTCAGAGCGATGGAA
GAGCTTTACGACGAAGGAAAACCGGTGGACGTGGTTTCCGTCTGTGACAA      200
GCTTCAAAGCATGGGAAAACTCGAGGAAGTAGGTGGAGATCTGGAAGTGG
CCCAGCTCGCTGAGGCTGTGCCCAGTTCTGCACACGCACTTCACTACGCG      300
GAGATCGTCAAGGAAAAATCCATTCTGAGGAAACTCATTGAGATCTCCAG
AAAAATCTCAGAAAGTGCCTACATGGAAGAAGATGTGGAGATCCTGCTCG      400
ACAACGCAGAAAAGATGATCTTCGAGATCTCAGAGATGAAAACGACAAAA
TCCTACGATCATCTGAGAGGCATCATGCACCGGGTGTTTGAAAACCTGGA      500
GAACTTCAGGGAAAGAGCCAACCTTATAGAACCCGGTGTGCTCATAACGG
GACTACCAACGGGATTCAAAAGTCTGGACAAACAGACCACAGGGTTCCAC      600
AGCTCCGATCTGGTGATAATAGCAGCGAGACCCTCCATGGGAAAAACCTC
CTTCGCACTCTCAATAGCGAGGAACATGGCTGTCAATTTCGAAATCCCCG      700
TCGGAATATTCAGTCTCGAGATGTCCAAGGAACAGCTCGCTCAAAGACTA
CTCAGCATGGAGTCCGGTGTGGATCTTTACAGCATCAGAACAGGATACCT      800
GGATCAGGAGAAGTGGGAAAGACTCACAATAGCGGCTTCTAAACTCTACA
AAGCACCCATAGTTGTGGACGATGAGTCACTCCTCGATCCGCGATCGTTG      900
AGGGCAAAAGCGAGAAGGATGAAAAAAGAATACGATGTAAAAGCCATTTT
TGTCGACTATCTCCAGCTCATGCACCTGAAAGGAAGAAAAGAAAGCAGAC     1000
AGCAGGAGATATCCGAGATCTCGAGATCTCTGAAGCTCCTTGCGAGGGAA
CTCGACATAGTGGTGATAGCGCTTTCACAGCTTTCGAGGGCCGTAGAACA     1100
GAGAGAAGACAAAAGACCGAGGCTGAGTGACCTCAGGGAATCCGGTGCGA
TAGAACAGGACGCAGACACAGTCATCTTCATCTACAGGGAGGAATATTAC     1200
AGGAGCAAAAAATCCAAAGAGGAAAGCAAGCTTCACGAACCTCACGAAGC
TGAAATCATAATAGGTAAACAGAGAAACGGTCCCGTTGGAACGATCACTC     1300
TGATCTTCGACCCCAGAACGGTTACGTTCCATGAAGTCGATGTGGTGCAT
TCA                                                    1353
```

FIG. 68

```
MRVPPHNLEAEVAVLGSILIDPSVINDVLEILSHEDFYLKKHQHIFRAME
ELYDEGKPVDVVSVCDKLQSMGKLEEVGGDLEVAQLAEAVPSSAHALHYA     100
EIVKEKSILRKLIEISRKISESAYMEEDVEILLDNAEKMIFEISEMKTTK
SYDHLRGIMHRVFENLENFRERANLIEPGVLITGLPTGFKSLDKQTTGFH     200
SSDLVIIAARPSMGKTSFALSIARNMAVNFEIPVGIFSLEMSKEQLAQRL
LSMESGVDLYSIRTGYLDQEKWERLTIAASKLYKAPIVVDDESLLDPRSL     300
RAKARRMKKEYDVKAIFVDYLQLMHLKGRKESRQQEISEISRSLKLLARE
LDIVVIALSQLSRAVEQREDKRPRLSDLRESGAIEQDADTVIFIYREEYY     400
RSKKSKEESKLHEPHEAEIIIGKQRNGPVGTITLIFDPRTVTFHEVDVVH
S                                                       451
```

FIG. 69

GTGATTCCTCGAGAGGTCATCGAGGAAATAAAAGAAAAGGTTGACATCGT
AGAGGTCATTTCCGAGTACGTGAATCTTACCCGGGTAGGTTCCTCCTACA 100
GGGCTCTCTGTCCCTTTCATTCAGAAACCAATCCTTCTTTCTACGTTCAT
CCGGGTTTGAAGATATACCATTGTTTCGGCTGCGGTGCGAGTGGAGACGT 200
CATCAAATTTCTTCAAGAAATGGAAGGGATCAGTTTCCAGGAAGCGCTGG
AAAGACTTGCCAAAAGAGCTGGGATTGATCTTTCTCTCTACAGAACAGAA 300
GGGACTTCTGAATACGGAAAATACATTCGTTTGTACGAAGAAACGTGGAA
AAGGTACGTCAAAGAGCTGGAGAAATCGAAAGAGGCAAAAGACTATTTAA 400
AAAGCAGAGGCTTCTCTGAAGAAGATATAGCAAAGTTCGGCTTTGGGTAC
GTCCCCAAGAGATCCAGCATCTCTATAGAAGTTGCAGAAGGCATGAACAT 500
AACACTGGAAGAACTTGTCAGATACGGTATCGCGCTGAAAAAGGGTGATC
GATTCGTTGATAGATTCGAAGGAAGAATCGTTGTTCCAATAAAGAACGAC 600
AGTGGTCATATTGTGGCTTTTGGTGGGCGTGCTCTCGGCAACGAAGAACC
GAAGTATTTGAACTCTCCAGAGACCAGGTATTTTTCGAAGAAGAAGACCC 700
TTTTTCTCTTCGATGAGGCGAAAAAAGTGGCAAAAGAGGTTGGTTTTTTC
GTCATCACCGAAGGCTACTTCGACGCGCTCGCATTCAGAAAGGATGGAAT 800
ACCAACGGCGGTCGCTGTTCTTGGGGCGAGTCTTTCAAGAGAGGCGATTC
TAAAACTTTCGGCGTATTCGAAAAACGTCATACTGTGTTTCGATAATGAC 900
AAAGCAGGCTTCAGAGCCACTCTCAAATCCCTCGAGGATCTCCTAGACTA
CGAATTCAACGTGCTTGTGGCAACCCCCTCTCCTTACAAAGACCCAGATG 1000
AACTCTTTCAGAAAGAAGGAGAAGGTTCATTGAAAAAGATGCTGAAAAAC
TCGCGTTCGTTCGAATATTTTCTGGTGACGGCTGGTGAGGTCTTCTTTGA 1100
CAGGAACAGCCCCGCGGGTGTGAGATCCTACCTTTCTTTCCTCAAAGGTT
GGGTCCAAAAGATGAGAAGGAAAGGATATTTGAAACACATAGAAAATCTC 1200
GTGAATGAGGTTTCATCTTCTCTCCAGATACCAGAAAACCAGATTTTGAA
CTTTTTTGAAAGCGACAGGTCTAACACTATGCCTGTTCATGAGACCAAGT 1300
CGTCAAAGGTTTACGATGAGGGGAGAGGACTGGCTTATTTGTTTTTGAAC
TACGAGGATTTGAGGGAAAAGATTCTGGAACTGGACTTAGAGGTACTGGA 1400
AGATAAAAACGCGAGGGAGTTTTTCAAGAGAGTCTCACTGGGAGAAGATT
TGAACAAAGTCATAGAAAACTTCCCAAAAGAGCTGAAAGACTGGATTTTT 1500
GAGACAATAGAAAGCATTCCTCCTCCAAAGGATCCCGAGAAATTCCTCGG
TGACCTCTCCGAAAAGTTGAAAATCCGACGGATAGAGAGACGTATCGCAG 1600
AAATAGATGATATGATAAAGAAAGCTTCAAACGATGAAGAAAGGCGTCTT
CTTCTCTCTATGAAAGTGGATCTCCTCAGAAAAATAAAGAGGAGG 1695

FIG. 70

```
MIPREVIEEIKEKVDIVEVISEYVNLTRVGSSYRALCPFHSETNPSFYVH
PGLKIYHCFGCGASGDVIKFLQEMEGISFQEALERLAKRAGIDLSLYRTE   100
GTSEYGKYIRLYEETWKRYVKELEKSKEAKDYLKSRGFSEEDIAKFGFGY
VPKRSSISIEVAEGMNITLEELVRYGIALKKGDRFVDRFEGRIVVPIKND   200
SGHIVAFGGRALGNEEPKYLNSPETRYFSKKKTLFLFDEAKKVAKEVGFF
VITEGYFDALAFRKDGIPTAVAVLGASLSREAILKLSAYSKNVILCFDND   300
KAGFRATLKSLEDLLDYEFNVLVATPSPYKDPDELFQKEGEGSLKKMLKN
SRSFEYFLVTAGEVFFDRNSPAGVRSYLSFLKGWVQKMRRKGYLKHIENL   400
VNEVSSSLQIPENQILNFFESDRSNTMPVHETKSSKVYDEGRGLAYLFLN
YEDLREKILELDLEVLEDKNAREFFKRVSLGEDLNKVIENFPKELKDWIF   500
ETIESIPPPKDPEKFLGDLSEKLKIRRIERRIAEIDDMIKKASNDEERRL
LLSMKVDLLRKIKRR                                      565
```

FIG. 71

```
ATGGCTCTACACCCGGCTCACCCTGGGGCAATAATCGGGCACGAGGCCGT
TCTCGCCCTCCTTCCCCGCCTCACCGCCCAGACCCTGCTCTTCTCCGGCC   100
CCGAGGGGGTGGGGCGGCGCACCGTGGCCCGCTGGTACGCCTGGGGGCTC
AACCGCGGCTTCCCCCCGCCCTCCCTGGGGGAGCACCCGGACGTCCTCGA   200
GGTGGGGCCCAAGGCCCGGGACCTCCGGGGCCGGGCCGAGGTGCGGCTGG
AGGAGGTGGCGCCCCTCTTGGAGTGGTGCTCCAGCCACCCCCGGGAGCGG   300
GTGAAGGTGGCCATCCTGGACTCGGCCCACCTCCTCACCGAGGCCGCCGC
CAACGCCCTCCTCAAGCTCCTGGAGGAGCCCCCTTCCTACGCCCGCATCG   400
TCCTCATCGCCCCAAGCCGCGCCACCCTCCTCCCCACCCTGGCCTCCCGG
GCCACGGAGGTGGCATTCGCCCCCGTGCCCGAGGAGGCCCTGCGCGCCCT   500
CACCCAGGACCCGGAGCTCCTCCGCTACGCCGCCGGGGCCCCGGGCCGCC
TCCTTAGGGCCCTCCAGGACCCGGAGGGGTACCGGGCCCGCATGGCCAGG   600
GCGCAAAGGGTCCTGAAAGCCCCGCCCCTGGAGCGCCTCGCTTTGCTTCG
GGAGCTTTTGGCCGAGGAGGAGGGGGTCCACGCCCTCCACGCCGTCCTAA   700
AGCGCCCGGAGCACCTCCTTGCCCTGGAGCGGGCGCGGGAGGCCCTGGAG
GGGTACGTGAGCCCCGAGCTGGTCCTCGCCCGGCTGGCCTTAGACTTAGA   800
GACA
```

FIG. 72

```
MALHPAHPGAIIGHEAVLALLPRLTAQTLLFSGPEGVGRRTVARWYAWGL
NRGFPPPSLGEHPDVLEVGPKARDLRGRAEVRLEEVAPLLEWCSSHPRER   100
VKVAILDSAHLLTEAAANALLKLLEEPPSYARIVLIAPSRATLLPTLASR
ATEVAFAPVPEEALRALTQDPELLRYAAGAPGRLLRALQDPEGYRARMAR   200
AQRVLKAPPLERLALLRELLAEEEGVHALHAVLKRPEHLLALERAREALE
GYVSPELVLARLALDLET                                   268
```

FIG. 73

```
ATGCTGGACCTGAGGGAGGTGGGGGAGGCGGAGTGGAAGGCCCTAAAGCC
CCTTTTGGAAAGCGTGCCCGAGGGCGTCCCCGTCCTCCTCCTGGACCCTA    100
AGCCAAGCCCCTCCCGGGCGGCCTTCTACCGGAACCGGGAAAGGCGGGAC
TTCCCCACCCCCAAGGGGAAGGACCTGGTGCGGCACCTGGAAAACCGGGC    200
CAAGCGCCTGGGGCTCAGGCTCCCGGGCGGGGTGGCCCAGTACCTGGCCT
CCCTGGAGGGGGACCTCGAGGCCCTGGAGCGGGAGCTGGAGAAGCTTGCC    300
CTCCTCTCCCCACCCCTCACCCTGGAGAAGGTGGAGAAGGTGGTGGCCCT
GAGGCCCCCCCTCACGGGCTTTGACCTGGTGCGCTCCGTCCTGGAGAAGG    400
ACCCCAAGGAGGCCCTCCTGCGCCTAGGCGGCCTCAAGGAGGAGGGGGAG
GAGCCCCTCAGGCTCCTCGGGGCCCTCTCCTGGCAGTTCGCCCTCCTCGC    500
CCGGGCCTTCTTCCTCCTCCGGGAAAACCCCAGGCCCAAGGAGGAGGACC
TCGCCCGCCTCGAGGCCCACCCCTACGCCGCCCGCCGCGCCCTGGAGGCG    600
GCGAAGCGCCTCACGGAAGAGGCCCTCAAGGAGGCCCTGGACGCCCTCAT
GGAGGCGGAAAAGAGGGCCAAGGGGGGGAAAGACCCGTGGCTCGCCCTGG    700
AGGCGGCGGTCCTCCGCCTCGCCCGTTGA
```

FIG. 74

```
MVIAFTGDPFLAREALLEEARLRGLSRFTEPTPEALAQALAPGLFGGGGA
MLDLREVGEAEWKALKPLLESVPEGVPVLLLDPKPSPSRAAFYRNRERRD    100
FPTPKGKDLVRHLENRAKRLGLRLPGGVAQYLASLEGDLEALERELEKLA
LLSPPLTLEKVEKVVALRPPLTGFDLVRSVLEKDPKEALLRLGGLKEEGE    200
EPLRLLGALSWQFALLARAFFLLRENPRPKEEDLARLEAHPYAARRALEA
AKRLTEEALKEALDALMEAEKRAKGGKDPWLALEAAVLRLAR            292
```

FIG. 75

```
ATGGCTCGAGGCCTGAACCGCGTTTTCCTCATCGGCGCCCTCGCCACCCG
GCCGGACATGCGCTACACCCCGGCGGGGCTCGCCATTTTGGACCTGACCC   100
TCGCCGGTCAGGACCTGCTTCTTTCCGATAACGGGGGGGAACCGGAGGTG
TCCTGGTACCACCGGGTGAGGCTCTTAGGCCGCCAGGCGGAGATGTGGGG   200
CGACCTCTTGGACCAAGGGCAGCTCGTCTTCGTGGAGGGCCGCCTGGAGT
ACCGCCAGTGGGAAAGGGAGGGGGAGAAGCGGAGCGAGCTCCAGATCCGG   300
GCCGACTTCCGGACCCCCTGGACGACCGGGGGAAGAAGCGGGCGGAGGAC
AGCCGGGGCCAGCCCAGGCTCCGCGCCGCCCTGAACCAGGTCTTCCTCAT   400
GGGCAACCTGACCCGGGACCCGGAACTCCGCTACACCCCCCAGGGCACCG
CGGTGGCCCGGCTGGGCCTGGCGGTGAACGAGCGCCGCCAGGGGGCGGAG   500
GAGCGCACCCACTTCGTGGAGGTTCAGGCCTGGCGCGACCTGGCGGAGTG
GGCCGCCGAGCTGAGGAAGGGCGACGGCCTTTTCGTGATCGGCAGGTTGG   600
TGAACGACTCCTGGACCAGCTCCAGCGGCGAGCGGCGCTTCCAGACCCGT
GTGGAGGCCCTCAGGCTGGAGCGCCCCACCCGTGGACCTGCCCAGGCCTG   700
CCCAGGCCGGCGGAACAGGTCCCGCGAAGTCCAGACGGGTGGGGTGGACA
TTGACGAAGGCTTGGAAGACTTTCCGCCGGAGGAGGATTTGCCGTTTTGA   800
GCACGAA
```

FIG. 76

```
MARGLNRVFLIGALATRPDMRYTPAGLAILDLTLAGQDLLLSDNGGEPEV
SWYHRVRLLGRQAEMWGDLLDQGQLVFVEGRLEYRQWEREGEKRSELQIR   100
ADFLDPLDDRGKKRAEDSRGQPRLRAALNQVFLMGNLTRDPELRYTPQGT
AVARLGLAVNERRQGAEERTHFVEVQAWRDLAEWAAELRKGDGLFVIGRL   200
VNDSWTSSSGERRFQTRVEALRLERPTRGPAQACPGRRNRSREVQTGGVD
IDEGLEDFPPEEDLPF                                     266
```

FIG. 77

```
AATTCCGACATTTCAATTGAATCGTTTATTCCGCTTGAAAAAGAAGGCAA
GTTGCTCGTTGATGTGAAAAGACCGGGGAGCATCGTACTGCAGGCGCGCT    100
TTTTCTCTGAAATCGTGAAAAAACTGCCGCAACAAACGGTGGAAATCGAA
ACGGAAGACAACTTTTTGACGATCATCCGCTCGGGGCACTCAGAATTCCG    200
CCTCAATGGGCTAAACGCCGACGAATATCCGCGCCTGCCGCAAATTGAAG
AAGAAAACGTGTTTCAAATCCCGGCTGATTTATTGAAAACCGTGATTCGG    300
CAAACGGTGTTCGCCGTTTCTACATCGGAAACGCGCCCAATCTTGACAGG
TGTCAACTGGAAAGTTGAACATGGCGAGCTTGTCTGCACAGCGACCGACA    400
GTCATCGCTTAGCCATGCGCAAAGTGAAAATTGAGTCGGAAAATGAAGTA
TCATACAACGTCGTCATCCCTGGAAAAAGTCTTAATGAGCTCAGCAAAAT    500
TTTGGATGACGGCAACCACCCGGTGGACATCGTCATGACAGCCAATCAAG
TGCTATTTAAGGCCGAGCACCTTCTCTTCTTTTCCCGGCTGCTTGACGGC    600
AACTATCCGGAGACGGCCCGCTTGATTCCAACAGAAAGCAAAACGACCAT
GATCGTCAATGCAAAAGAGTTTCTGCAGGCAATCGACCGAGCGTCCTTGC    700
TTGCTCGAGAAGGAAGGAACAACGTTGTGAAACTGACGACGCTTCCTGGA
GGAATGCTCGAAATTTCTTCGATTTCTCCGAGATCGGGAAAGTGACGGAG    800
CAGCTGCAAACGGAGTCTCTTGAAGGGGAAGAGTTGAACATTTCGTTCAG
CGCGAAATATATGATGGACGCGTTGCGGGCGCTTGATGGAACAGACATTT    900
CAAATCAGCTTCACTGGGGCCATGCGGCCGTTCCTGTTGCGCCCGCTTCA
ACCGATTCGATGCTTCAGCTCATTTTGCCGGTGAGAACATAT            992
```

FIG. 78

```
NSDISIIESFIPLEKEGKLLVDVKRPGSIVLQARFFSEIVKKLPQQTVEI
ETEDNFLTIIRSGHSEFRLNGLNADEYPRLPQIEEENVFQIPADLLKTVI    100
RQTVFAVSTSETRPILTGVNWKVEHGELVCTATDSHRLAMRKVKIIESEN
EVSYNVVIPGKSLNELSKIILDDGNHPVDIVMTANQVLFKAEHLLFFSRL    200
LDGNYPETARLIPTESKTTMIVNAKEFLQAIDRASLLAREGRNNVVKLTT
LPGGMLEISSISPEIGKVTEQLQTESLEGEELNISFSAKYMMDALRALDG    300
TDIQISFTGAMRPFLLRPLHTDSMLQLILPVRTY
```

FIG. 79

ATGATTAACCGCGTCATTTTGGTCGGCAGGTTAACGAGAGATCCGGAGTT
GCGTTACACTCCAAGCGGAGTGGCTGTTGCCACGTTTACGCTCGCGGTCA 100
ACCGTCCGTTTACAAATCAGCAGGGCGAGCGGGAAACGGATTTTATTCAA
TGTGTCGTTTGGCGCCGCCAGGCGGAAAACGTCGCCAACTTTTTGAAAAA 200
GGGGAGCTTGGCTGGTGTCGATGGCCGACTGCAAACCCGCAGCTATGAAA
ATCAAGAAGGTCGGCGTGTGTACGTGACGGAAGTGGTGGCTGATAGCGTC 300
CAATTTCTTGAGCCGAAAGGAACGAGCGAGCAGCGAGGGGCGACAGCAGG
CGGCTACTATGGGGATCCATTCCCATTCGGGCAAGATCAGAACCACCAAT 400
ATCCGAACGAAAAAGGGTTTGGCCGCATCGATGACGATCCTTTCGCCAAT
GACGGCCAGCCGATCGATATTTCTGATGATGATTTGCCGTTT 492

FIG. 80

MINRVILVGRLTRDPELRYTPSGVAVATFTLAVNRPFTNQSYENQEGRRV
YVTEVVADSVQFLEPKGTSEQRGATAGGYYQGERETDFIQCVVWRRQAEN 100
VANFLKKGSLAGVDGRLQTRGDPFPFGQDQNHQYPNEKGFGRIDDDPFAN
DGQPIDISDDDLPF 164

FIG. 81

ATGCTGGAACGCGTATGGGGAAACATTGAAAAACGGCGTTTTTCTCCCCT
TTATTTATTATACGGCAATGAGCCGTTTTTATTAACGGAAACGTATGAGC 100
GATTGGTGAACGCAGCGCTTGGCCCCGAGGAGCGGGAGTGGAACTTGGCT
GTGTACGACTGCGAGGAAACGCCGATCGAGGCGGCGCTTGAGGAGGCCGA 200
GACGGTGCCGTTTTTCGGCGAGCGGCGTGTCATTCTCATCAAGCATCCAT
ATTTTTTTACGTCTGAAAAAGAGAAGGAGATCGAACATGATTTGGCGAAG 300
CTGGAGGCGTACTTGAAGGCGCCGTCGCCGTTTTCGATCGTCGTCTTTTT
CGCGCCGTACGAGAAGCTTGATGAGCGAAAAAAAATTACGAAGCTCGCCA 400
AAGAGCAAAGCGAAGTCGTCATCGCCGCCCCGCTCGCCGAAGCGGAGCTG
CGTGCCTGGGTGCGGCGCCGCATCGAGAGCCAAGGGGCGCAAGCAAGCGA 500
CGAGGCGATTGATGTCCTGTTGCGGCGGGCCGGGACGCAGCTTTCCGCCT
TGGCGAATGAAATCGATAAATTGGCCCTGTTTGCCGGATCGGGCGGAACC 600
ATCGAGGCGGCGGCGGTTGAGCGGCTTGTCGCCCGCACGCCGGAAGAAAA
CGTATTTGTGCTTGTCGAGCAAGTGGCGAAGCGCGACATTCCAGCAGCGT 700
TGCAGACGTTTTATGATCTGCTTGAAAACAATGAAGAGCCGATCAAAATT
TTGGCGTTGCTCGCCGCCCATTTCCGCTTGCTTTCGCAAGTGAAATGGCT 800
TGCCTCCTTAGGCTACGGACAGGCGCAAATTGCTGCGGCGCTCAAGGTGC
ACCCGTTCCGCGTCAAGCTCGCTCTTGCTCAAGCGGCCCGCTTCGCTGAC 900
GGAGAGCTTGCTGAGGCGATCAACGAGCTCGCTGACGCCGATTACGAAGT
GAAAAGCGGGGCGGTCGATCGCCGGTTGGCCGTTGAGCTGCTTCTGATGC 1000
GCTGGGGCGCCCGCCCGGCGCAAGCGGGGCGCCACGGCCGGCGG

FIG. 82

MLERVWGNIEKRRFSPLYLLYGNEPFLLTETYERLVNAALGPEEREWNLA
VYDCEETPIEAALEEAETVPFFGERRVILIKHPYFFTSEKEKEIEHDLAK 100
LEAYLKAPSPFSIVVFFAPYEKLDERKKITKLAKEQSEVVIAAPLAEAEL
RAWVRRRIESQGAQASDEAIDVLLRRAGTQLSALANEIDKLALFAGSGGT 200
IEAAAVERLVARTPEENVFVLVEQVAKRDIPAALQTFYDLLENNEEPIKI
LALLAAHFRLLSQVKWLASLGYGQAQIAAALKVHPFRVKLALAQAARFAD 300
GELAEAINELADADYEVKSGAVDRRLAVELLLMRWGARPAQAGRHGRR

FIG. 83

ATGCGATGGGAACAGCTAGCGAAACGCCAGCCGGTGGTGGCGAAAATGCT
GCAAAGCGGCTTGGAAAAAGGGCGGATTTCTCATGCGTACTTGTTTGAGG 100
GGCAGCGGGGGACGGGCAAAAAAGCGGCCAGTTTGTTGTTGGCGAAACGT
TTGTTTTGTCTGTCCCCAATCGGAGTTTCCCCGTGTCTAGAGTGCCGCAA 200
CTGCCGGCGCATCGACTCCGGCAACCACCCTGACGTCCGGGTGATCGGCC
CAGATGGAGGATCAATCAAAAAGGAACAAATCGAATGGCTGCAGCAAGAG 300
TTCTCGAAAACAGCGGTCGAGTCGGATAAAAAAATGTACATCGTTGAGCA
CGCCGATCAAATGACGACAAGCGCTGCCAACAGCCTTCTGAAATTTTTGG 400
AAGAGCCGCATCCGGGGACGGTGGCGGTATTGCTGACTGAGCAATACCAC
CGCCTGCTAGGGACGATCGTTTCCCGCTGTCAAGTGCTTTCGTTCCGGCC 500
GTTGCCGCCGGCAGAGCTCGCCCAGGGACTTGTCGAGGAGCACGTGCCGT
TGCCGTTGGCGCTGTTGGCTGCCCATTTGACAAACAGCTTCGAGGAAGCA 600
CTGGCGCTTGCCAAAGATAGTTGGTTTGCCGAGGCGCGAACATTAGTGCT
ACAATGGTATGAGATGCTGGGCAAGCCGGAGCTGCAGCTTTTGTTTTTCA 700
TCCACGACCGCTTGTTTCCGCATTTTTTGGAAAGCCATCAGCTTGACCTT
GGACTTG 757

FIG. 84

MRWEQLAKRQPVVAKMLQSGLEKGRISHAYLFEGQRGTGKKAASLLLAKR
LFCLSPIGVSPCLECRNCRRIDSGNHPDVRVIGPDGGSIKKEQIEWLQQE 100
FSKTAVESDKKMYIVEHADQMTTSAANSLLKFLEEPHPGTVAVLLTEQYH
RLLGTIVSRCQVLSFRPLPPAELAQGLVEEHVPLPLALLAAHLTNSFEEA 200
LALAKDSWFAEARTLVLQWYEMLGKPELQLLFFIHDRLFPHFLESHQLDL
GL 252

FIG. 85

GTGGCATACCAAGCGTTATATCGCGTGTTTCGGCCGCAGCGCTTTGCGGA
CATGGTCGGCCAAGAACACGTGACCAAGACGTTGCAAAGCGCCCTGCTTC 100
AACATAAAATATCGCACGCTTACTTATTTTCCGGCCCGCGCGGTACAGGA
AAAACGAGCGCAGCGAAAATTTTCGCCAAGGCGGTCAACTGTGAACAGGC 200
GCCAGCGGCGGAGCCATGCAATGAGTGTCCAGCTTGCCTCGGCATTACGA
ATGGAACGGTTCCCGATGTGCTGGAAATTGACGCTGCTTCCAACAACCGC 300
GTCGATGAAATTCGTGATATCCGTGAGAAGGTGAAATTTGCGCCAACGTC
GGCCCGCTACAAAGTGTATATCATCGACGAGGTGCATATGCTGTCGATCG 400
GTGCGTTTAACGCGCTGTTGAAAACGTTGGAGGAGCCGCCGAAACACGTC
ATTTTCATTTTGGCCACGACCGAGCCGCACAAAATTCCGGCGACGATCAT 500
TTCCCGCTGCCAACGGTTCGATTTTCGCCGCATCCCGCTTCAGGCGATCG
TTTCACGGCTAAAGTACGTCGCAAGCGCCCAAGGTGTCGAGGCGTCAGAT 600
GAGGCATTGTCCGCCATCGCCCGTGCTGCAGACGGGGGGATGCGCGATGC
GCTCAGCTTGCTTGATCAAGCCATTTCGTTCAGCGACGGGAAACTTCGGC 700
TCGACGACGTGCTGGCGATGACCGGGGCTGCATCATTTGCCGCCTTATCG
AGCTTCATCGAAGCCATCCACCGCAAAGATACAGCGGCGGTTCTTCAGCA 800
CTTGGAAACGATGATGGCGCAAGGGAAAGATCCGCATCGTTTGGTTGAAG
ACTTGATTTTGTACTATCGCGATTTATTGCTGTACAAAACCGCTCCCTAT 900
GTGGAGGGAGCGATTCAAATTGCTGTCGTTGACGAAGCGTTCACTTCACT
GTCGGAAATGATTCCGGTTTCCAATTTATACGAGGCCATCGAGTTGCTGA 1000
ACAAAAGCCAGCAAGAGATGAAGTGGACAAACCACCCGCGCCTTCTGTTG
GAAGTGGCGCTTGTGAAACTTTGCCATCCATCAGCCGCCGCCCCGTCGCT 1100
GTCGGCTTCCGAGTTGGAACCGTTGATAAAGCGGATTGAAACGCTGGAGG
CGGAATTGCGGCGCCTGAAGGAACAACCGCCTGCCCCTCCGTCGACCGCC 1200
GCGCCGGTGAAAAAACTGTCCAAACCGATGAAAACGGGGGGGATATAAAGC
CCCGGTTGGCCGCATTTACGAGCTGTTGAAACAGGCGACGCATGAAGATT 1300
TAGCTTTGGTGAAAGGATGCTGGGCGGATGTGCTCGACACGTTGAAACGG
CAGCATAAAGTGTCGCACGCTGCCTTGCTGCAAGAGAGCGAGCCGGTTGC 1400
AGCGAGCGCCTCAGCGTTTGTATTAAAATTCAAATACGAAATCCACTGCA
AAATGGCGACCGATCCCACAAGTTCGGTCAAAGAAAACGTCGAAGCGATT 1500
TTGTTTGAGCTGACAAACCGCCGCTTTGAAATGGTAGCCATTCCGGAGGG
AGAATGGGGAAAAATAAGAGAAGAGTTCATCCGCAATAAGGACGCCATGG 1600
TGGAAAAAAGCGAAGAAGATCCGTTAATCGCCGAAGCGAAGCGGCTGTTT
GGCGAAGAGCTGATCGAAATTAAAGAA 1677

FIG. 86

VAYQALYRVFRPQRFADMVGQEHVTKTLQSALLQHKISHAYLFSGPRGTG
KTSAAKIFAKAVNCEQAPAAEPCNECPACLGITNGTVPDVLEIDAASNNR 100
VDEIRDIREKVKFAPTSARYKVYIIDEVHMLSIGAFNALLKTLEEPPKHV
IFILATTEPHKIPATIISRCQRFDFRRIPLQAIVSRLKYVASAQGVEASD 200
EALSAIARAADGGMRDALSLLDQAISFSDGKLRLDDVLAMTGAASFAALS
SFIEAIHRKDTAAVLQHLETMMAQGKDPHRLVEDLILYYRDLLLYKTAPY 300
VEGAIQIAVVDEAFTSLSEMIPVSNLYEAIELLNKSQQEMKWTNHPRLLL
EVALVKLCHPSAAAPSLSASELEPLIKRIETLEAELRRLKEQPPAPPSTA 400
APVKKLSKPMKTGGYKAPVGRIYELLKQATHEDLALVKGCWADVLDTLKR
QHKVSHAALLQESEPVAASASAFVLKFKYEIHCKMATDPTSSVKENVEAI 500
LFELTNRRFEMVAIPEGEWGKIREEFIRNKDAMVEKSEEDPLIAEAKRLF
GEELIEIKE 559

FIG. 87

ATGGTGACAAAAGAGCAAAAAGAGCGGTTTCTCATCCTGCTTGAGCAGCT
GAAGATGACGTCGGACGAATGGATGCCGCATTTTCGTGAGGCAGCCATTC 100
GCAAAGTCGTGATCGATAAAGAGGAGAAAAGCTGGCATTTTTATTTTCAG
TTCGACAACGTGCTGCCGGTTCATGTATACAAAACGTTTGCCGATCGGCT 200
GCAGACGGCGTTCCGCCATATCGCCGCCGTCCGCCATACGATGGAGGTCG
AAGCGCCGCGCGTAACTGAGGCGGATGTGCAGGCGTATTGGCCGCTTTGC 300
CTTGCCGAGCTGCAAGAAGGCATGTCGCCGCTTGTCGATTGGCTCAGCCG
GCAGACGCCTGAGCTGAAAGGAAACAAGCTGCTTGTCGTTGCCCGCCATG 400
AAGCGGAAGCGCTGGCGATCAAACGGCGGTTCGCCAAAAAAATCGCTGAT
GTGTACGCTTCGTTTGGGTTTCCCCCCCTTCAGCTTGACGTCAGCGTCGA 500
GCCGTCCAAGCAAGAAATGGAACAGTTTTTGGCGCAAAAACAGCAAGAGG
ACGAAGAGCGAGCGCTTGCTGTACTGACCGATTTAGCGAGGGAAGAAGAA 600
AAGGCCGCGTCTGCGCCGCCGTCCGGTCCGCTTGTCATCGGCTATCCGAT
CCGCGACGAGGAGCCGGTGCGGCGGCTTGAAACGATCGTCGAAGAAGAGC 700
GGCGCGTCGTTGTGCAAGGCTATGTATTTGACGCCGAAGTGAGCGAATTA
AAAAGCGGCCGCACGCTGTTGACCATGAAAATCACAGATTACACGAACTC 800
GATTTTAGTCAAAATGTTCTCGCGCGACAAAGAGGACGCCGAGCTTATGA
GCGGCGTCAAAAAAGGCATGTGGGTGAAAGTGCGCGGCAGCGTGCAAAAC 900
GATACGTTCGTCCGTGATTTGGTCATCATCGCCAACGATTTGAACGAAAT
CGCCGCAAACGAACGGCAAGATACGGCGCCGGAAGGGGAAAAGAGGGTCG 1000
AGCTCCATTTGCATACCCCGATGAGCCAAATGGACGCGGTCACCTCGGTG
ACAAAACTCATTGAGCAAGCGAAAAAATGGGGGCATCCGGCGATCGCCGT 1100
CACCGACCATGCCGTTGTTCAGTCGTTTCCGGAGGCCTACAGCGCGGCGA
AAAAACACGGCATGAAGGTCATTTACGGCCTTGAGGCGAACATCGTCGAC 1200
GATGGCGTGCCGATCGCCTACAATGAGACGCACCGCCGTCTTTCGGAGGA
AACGTACGTCGTCTTTGACGTCGAGACGACGGGCCTGTCGGCTGTGTACA 1300
ATACGATCATTGAGCTGGCGGCGGTGAAAGTGAAAGACGGCGAGATCATC
GACCGATTCATGTCGTTTGCCAACCCTGGACATCCGTTGTCGGTGACAAC 1400
GATGGAGCTGACTGGGATCACCGATGAGATGGTGAAAGACGCCCCGAAGC
CGGACGAGGTGCTAGCCCGTTTTGTTGACTGGGCCGGCGATGCGACGCTT 1500
GTTGCCCACAACGCCAGCTTTGACATCGGTTTTTTAAACGCGGGCCTCGC
TCGCATGGGGCGCGGCAAAATCGCGAATCCAGTCATCGATACGCTCGAGC 1600
TGGCCCGTTTTTTATACCCGGATTTGAAAAACCATCGGCTCAATACATTG
TGCAAAAAATTTGACATTGAATTGACGCAGCATCACCGCGCCATCTACGA 1700
CGCGGAGGCGACCGGGCATTTGCTTATGCGGCTGTTGAAGGAAGCGGAAG
AGCGCGGCATACTGTTTCATGACGAATTAAACAGCCGCACGCACAGCGAA 1800
GCGTCCTATCGGCTTGCGCGCCCGTTCCATGTGACGCTGTTGGCGCAAAA
CGAGACTGGATTGAAAAATTTGTTCAAGCTTGTGTCATTGTCGCACATTC 1900
AATATTTTCACCGTGTGCCGCGCATCCCGCGCTCCGTGCTCGTCAAGCAC
CGCGACGGCCTGCTTGTCGGCTCGGGCTGCGACAAAGGAGAGCTGTTTGA 2000
CAACTTGATCCAAAAGGCGCCGGAAGAAGTCGAAGACATCGCCCGTTTTT
ACGATTTTCTTGAAGTGCATCCGCCGGACGTGTACAAGCCGCTCATCGAG 2100
ATGGATTATGTGAAAGACGAAGAGATGATCAAAAACATCATCCGCAGCAT
CGTCGCCCTTGGTGAGAAGCTTGACATCCCGGTTGTCGCCACTGGCAACG 2200

FIG. 88A

TCCATTACTTGAACCCAGAAGATAAAATTTACCGGAAAATCTTAATCCAT
TCGCAAGGCGGGGCGAATCCGCTCAACCGCCATGAACTGCCGGATGTATA 2300
TTTCCGTACGACGAATGAAATGCTTGACTGCTTCTCGTTTTTAGGGCCGG
AAAAAGCGAAGGAAATCGTCGTTGACAACACGCAAAAAATCGCTTCGTTA 2400
ATCGGCGATGTCAAGCCGATCAAAGATGAGCTGTATACGCCGCGCATTGA
AGGGGCGGACGAGGAAATCAGGGAAATGAGCTACCGGCGGGCGAAGGAAA 2500
TTTACGGCGACCCGTTGCCGAAACTTGTTGAAGAGCGGCTTGAGAAGGAG
CTAAAAAGCATCATCGGCCATGGCTTTGCCGTCATTTATTTGATCTCGCA 2600
CAAGCTTGTGAAAAAATCGCTCGATGACGGCTACCTTGTCGGGTCGCGCG
GATCGGTCGGCTCGTCGTTTGTCGCGACGATGACGGAAATCACCGAGGTC 2700
AATCCGCTGCCGCCGCATTACGTTTGCCCGAACTGCAAGCATTCGGAGTT
CTTTAACGACGGTTCAGTCGGCTCAGGGTTTGATTTGCCGGATAAAAACT 2800
GCCCGCGATGTGGGACGAAATACAAGAAAGACGGGCACGACATCCCGTTT
GAGACGTTTCTCGGCTTTAAAGGCGACAAAGTGCCGGATATCGACTTGAA 2900
CTTTTCCGGCGAATACCAGCCGCGCGCCCACAACTATACGAAAGTGCTGT
TTGGCGAAGACAACGTCTACCGCGCCGGGACGATTGGCACGGTCGCTGAC 3000
AAAACGGCGTACGGATTTGTCAAAGCGTATGCGAGCGACCATAACTTAGA
GCTGCGCGGCGCGGAAATCGACGGCTCGCGGCTGGCTGCACCGGGGTGAA 3100
GCGGACGACCGGGCAGCATCCGGGCGGCATCATCGTCGTCCCGGATTATA
TGGAAATTTACGATTTTACGCCGATTCAATATCCGGCCGATGACACGTCC 3200
TCTGAATGGCGGACGACCCATTTCGACTTCCATTCGATCCACGACAATTT
GTTGAAGCTCGATATTCTCGGGCACGACGATCCGACGGTCATTCGCATGC 3300
TGCAAGATTTAAGCGGCATCGATCCGAAAACGATCCCGACCGACGACCCG
GATGTGATGGGCATTTTCAGCAGCACCGAGCCGCTTGGCGTTACGCCGGA 3400
GCAAATCATGTGCAATGTCGGCACGATCGGCATTCCGGAGTTTGGCACGC
GCTTCGTTCGGCAAATGTTGGAAGAGACAAGGCCAAAAACGTTTTCCGAA 3500
CTCGTGCAAATTTCCGGCTTGTCGCACGGCACCGATGTGTGGCTCGGCAA
CGCGCAAGAGCTCATTCAAAACGGCACGTGTACGTTATCGGAAGTCATCG 3600
GCTGCCGCGACGACATTATGGTCTATTTGATTTACCGCGGGCTCGAGCCG
TCGCTCGCTTTTAAAATCATGGAATCCGTGCGCAAAGGAAAAGGCTTAAC 3700
GCCGGAGTTTGAAGCAGAAATGCGCAAACATGACGTGCCGGAGTGGTACA
TCGATTCATGCAAAAAAATCAAGTACATGTTCCCGAAAGCGCACGCCGCC 3800
GCCTACGTGTTAATGGCGGTGCGCATCGCCTACTTTAAGGTGCACCATCC
GCTTTTGTATTACGCGTCGTACTTTACGGTGCGGGCGGAGGACTTTGACC 3900
TTGACGCCATGATCAAAGGATCACCCGCCATTCGCAAGCGGATTGAGGAA
ATCAACGCCAAAGGCATTCAGGCGACGGCGAAAGAAAAAAGCTTGCTCAC 4000
GGTTCTTGAGGTGGCCTTAGAGATGTGCGAGCGCGGCTTTTCCTTTAAAA
ATATCGATTTGTACCGCTCGCAGGCGACGGAATTCGTCATTGACGGCAAT 4100
TCTCTCATTCCGCCGTTCAACGCCATTCCGGGGCTTGGGACGAACGTGGC
GCAGGCGATCGTGCGCGCCCGCGAGGAAGGCGAGTTTTTGTCGAAGGAGG 4200
ATTTGCAACAGCGCGGCAAATTGTCGAAAACGCTGCTCGAGTATCTAGAA
AGCCGCGGCTGCCTTGACTCGCTTCCAGACCATAACCAGCTGTCGCTGTT 4300
T

FIG. 88B

MVTKEQKERFLILLEQLKMTSDEWMPHFREAAIRKVVIDKEEKSWHFYFQ
FDNVLPVHVYKTFADRLQTAFRHIAAVRHTMEVEAPRVTEADVQAYWPLC 100
LAELQEGMSPLVDWLSRQTPELKGNKLLVVARHEAEALAIKRRFAKKIAD
VYASFGFPPLQLDVSVEPSKQEMEQFLAQKQQEDEERALAVLTDLAREEE 200
KAASAPPSGPLVIGYPIRDEEPVRRLETIVEEERRVVVQGYVFDAEVSEL
KSGRTLLTMKITDYTNSILVKMFSRDKEDAELMSGVKKGMWVKVRGSVQN 300
DTFVRDLVIIANDLNEIAANERQDTAPEGEKRVELHLHTPMSQMDAVTSV
TKLIEQAKKWGHPAIAVTDHAVVQSFPEAYSAAKKHGMKVIYGLEANIVD 400
DGVPIAYNETHRRLSEETYVVFDVETTGLSAVYNTIIELAAVKVKDGEII
DRFMSFANPGHPLSVTTMELTGITDEMVKDAPKPDEVLARFVDWAGDATL 500
VAHNASFDIGFLNAGLARMGRGKIANPVIDTLELARFLYPDLKNHRLNTL
CKKFDIELTQHHRAIYDAEATGHLLMRLLKEAEERGILFHDELNSRTHSE 600
ASYRLARPFHVTLLAQNETGLKNLFKLVSLSHIQYFHRVPRIPRSVLVKH
RDGLLVGSGCDKGELFDNLIQKAPEEVEDIARFYDFLEVHPPDVYKPLIE 700
MDYVKDEEMIKNIIRSIVALGEKLDIPVVATGNVHYLNPEDKIYRKILIH
SQGGANPLNRHELPDVYFRTTNEMLDCFSFLGPEKAKEIVVDNTQKIASL 800
IGDVKPIKDELYTPRIEGADEEIREMSYRRAKEIYGDPLPKLVEERLEKE
LKSIIGHGFAVIYLISHKLVKKSLDDGYLVGSRGSVGSSFVATMTEITEV 900
NPLPPHYVCPNCKHSEFFNDGSVGSGFDLPDKNCPRCGTKYKKDGHDIPF
ETFLGFKGDKVPDIDLNFSGEYQPRAHNYTKVLFGEDNVYRAGTIGTVAD 1000
KTAYGFVKAYASDHNLELRGAEIDLAAGCTGVKRTTGQHPGGIIVVPDYM
EIYDFTPIQYPADDTSSEWRTTHFDFHSIHDNLLKLDILGHDDPTVIRML 1100
QDLSGIDPKTIPTDDPDVMGIFSSTEPLGVTPEQIMCNVGTIGIPEFGTR
FVRQMLEETRPKTFSELVQISGLSHGTDVWLGNAQELIQNGTCTLSEVIG 1200
CRDDIMVYLIYRGLEPSLAFKIMESVRKGKGLTPEFEAEMRKHDVPEWYI
DSCKKIKYMFPKAHAAAYVLMAVRIAYFKVHHPLLYYASYFTVRAEDFDL 1300
DAMIKGSPAIRKRIEEINAKGIQATAKEKSLLTVLEVALEMCERGFSFKN
IDLYRSQATEFVIDGNSLIPPFNAIPGLGTNVAQAIVRAREEGEFLSKED 1400
LQQRGKLSKTLLEYLESRGCLDSLPDHNQLSLF

FIG. 89